FOURTH DRAFT

MODELING PROTOCOL

**Development of a
Photochemical Modeling System to
Support The Bay Area Air Quality Management District's
2004 State Implementation Plan**

Prepared for
Bay Area Air Quality Management District
939 Ellis Street
San Francisco, CA 94109

Prepared by

ENVIRON International Corporation
101 Rowland Way, Suite 220
Novato, CA 94945

ATMET, LLC
P.O. Box 19195
Boulder, CO 80308

Alpine Geophysics, LLC
7691 Alpine Rd.
La Honda, CA 94020

December 2002

TABLE OF CONTENTS

APPENDICES

Appendix A: EMS-95 Emissions Dataset Formats

TABLES

FIGURES

1. INTRODUCTION

The Bay Area Air Quality Management District (BAAQMD or District) is required to submit a State Implementation Plan (SIP) revision for ozone air quality to the California Air Resources Board (CARB) and to the U.S. Environmental Protection Agency (EPA) in 2004. The updated plan must show that the San Francisco Bay Area (Bay Area or SFBA) attains the 1-hour ozone National Ambient Air Quality Standards (NAAQS) by 2006. The forecast of ambient ozone air quality into the future is achieved through computer modeling.

This document establishes and describes the procedures that will be used to develop a new ozone modeling system and database for the San Francisco Bay Area. A Modeling Protocol such as this is essential whenever ozone modeling is carried out for the purpose of developing emission reduction strategies that may be included in a State Implementation Plan. The requirements for a Modeling Protocol are described in two guidance documents:

> "Guideline for Regulatory Application of the Urban Airshed Model," EPA-450/4-91-013, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC 27711. July 1991.

> "Technical Guidance Document: Photochemical Modeling," California Environmental Protection Agency, Air Resources Board. April 1992.

Other more recent modeling guidance includes a focus on how to use model results for attainment demonstrations, and newer methodologies established for modeling 8-hour ozone:

> "Guidance on Use of Modeled Results to Demonstrate Attainment of the Ozone NAAQS," EPA-454/B-95-007, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC 27711. 1996.

> "Draft Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-Hour Ozone NAAQS." EPA-454/R-99-004, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, NC 27711. May 1999.

The Modeling Protocol delineates the objectives, procedures, and expected results of the modeling study and sets up a process for participation between the regulators and stakeholders to avoid potential technical conflicts. Protocol development should be a dynamic process that is modified as new information is acquired. Proposed changes will be reviewed by interested parties and incorporated if approved. The California Clean Air Act (CCAA) requires that the Modeling Protocol be approved by both the local District and the CARB. The guidance suggests that at least two review groups be established to review and approve the Modeling Protocol and to review the results of the study as they become available. For the Bay Area, the District Board's executive committee will serve as the policy review group, and a Modeling Advisory Committee (MAC), including stakeholders and representatives from other agencies, has been assembled to review the technical aspects of the project.

This Modeling Protocol should be viewed as a "living" document. That is, after the District and MAC reviewed the initial draft, comments were compiled and responses developed. The Modeling Protocol will then continue to be revised as necessary and made available to the project participants. If new information necessitates updates to the modeling approach well into the study, the Modeling Protocol will be further revised to reflect the alternative methodology, and resubmitted to the participants for review. The development of the Modeling Protocol is viewed as a joint effort between the contracting Team, the District and the MAC. This Modeling Protocol identifies work being carried out by the CARB to prevent duplication of efforts, and emphasizes the synergy between the District's and CARB's modeling of the Central California Ozone Study (CCOS) episodes and modeling domains.

The remainder of this section provides some background information relevant to the ozone air quality problem in the SFBA, establishes the purpose and objectives of the current study along with an overview of the approach and relationship to the CCOS, provides a schedule for completion of the various tasks and associated deliverables, and lists the study participants. Section 2 discusses the episode selection; Section 3 provides a justification for model selection; and Sections 4 through 8 describe the meteorological modeling approach, emissions modeling approach, photochemical model input data preparation, base year photochemical model performance evaluation, and attainment year modeling.

BACKGROUND

The SFBA was initially classified as a "Moderate" nonattainment area for 1-hour ozone following the passage of the 1990 Clean Air Act Amendments (CAAA). The EPA approved the subsequent SIP for the area and found the area in attainment for ozone in May 1995. However, new violations occurred during the summer of 1995, prompting the EPA to reverse its finding in July 1998 and to declare the SFBA as nonattainment. Since this action occurred under Section 172 of the CAAA, the SFBA is now classified as "Other". In March 2001, the EPA disapproved portions of the SFBA SIP. As a result, the BAAQMD was required to submit a new SIP, which it did on November 30, 2001. Included in this SIP submittal is a commitment to revise this 2001 SIP by April 15, 2004 to demonstrate attainment of the 1-hour ozone standard by 2006.

PURPOSE AND OBJECTIVES OF THIS STUDY

The overall purpose for this study is divided into two distinct goals:

> Immediate and foremost goal:
> Provide the District with technical analysis and photochemical modeling results in support of the 2004 Bay Area SIP revision, to include assessment of projected future year ozone levels in the SFBA, examination of the control strategy effectiveness, and analyses of the impact of those strategies on regional ozone throughout central California.

Longer-range goal:
Provide the District with a modern tool base that they can use to build a modeling "climatology", consisting of many additional historical episodes with which to evaluate local/regional ozone patterns and issues surrounding inter-basin transport.

The District originally identified nine specific contract objectives that must be met for the successful execution of the project. They are:

1. All technical modeling development work is to be completed by July 1, 2003, while all future year analyses, SIP documentation, and transfer/training is to be completed by September 2003;
2. The modeling system will meet or exceed the requirements in the Request for Proposals;
3. A new computing platform will be acquired that will meet the requirements of the District staff;
4. All data and software will be installed and functional on the specified computing platform;
5. Meteorological and photochemical simulations will be provided for two CCOS episodes (a third episode in July 1999 has since been added):
   14-15 June, 2000
   30 July – 2 August, 2000
6. Emission inputs and scenario development will be generated using the CARB's EMS-95 setup and input databases;
7. Photochemical modeling will meet or exceed EPA/CARB model performance criteria;
8. Full documentation of all results will be provided as a Technical Support Document to the 2004 ozone SIP;
9. The modeling system and all results will be transferred to the District and personnel will be trained on all aspects of operation and evaluation.

These nine objectives will be met by completing the work in nine individual project task elements, as established by ENVIRON and the District:

TASK 1: Develop Modeling Protocol
TASK 2: Acquire Computing System
TASK 3: Simulate Meteorological Conditions for the 2000 Episodes
TASK 4: Produce Emission Inputs for Years 2000 and 2006
TASK 5: Simulate Base Year Ozone and Evaluate Photochemical Model Performance
TASK 6: Simulate Year 2006 Ozone and Examine Model Sensitivity to Emissions
TASK 7: Simulate Year 2006 SIP Control Strategies and Demonstrate Attainment
TASK 8: Management and Reporting
TASK 9: Transfer Project Computer and Modeling System with Training

Although this project is sponsored by the District, we recognize the importance of working closely with the CARB. We plan to fully engage the CARB technical staff in this effort, thus exploiting the relevant data, methods, and technical expertise at that agency. We have been assured by the CARB technical staff and management that they are enthusiastic to work with us to share CCOS data, methods, and results.

Overview of Approach

The EMS-95/RAMS/CAMx modeling system will be used for this study because it contains all of the technical features necessary to simulate ozone air quality in the SFBA. We believe that this system of models has the highest likelihood of generating SIP-quality photochemical modeling databases for the Bay Area.

The same EMS-95 emissions processor and input databases as used by the CARB will be used in this project to assure CARB compatibility and acceptability. We believe that this is an essential element of the study. Portions of the emissions database may not be ready from CARB in time to meet the schedule demands of this project, so we will maintain close contact with CARB staff so that we may identify which portions of the inventory we may need to process independently. ENVIRON's subcontractor Alpine Geophysics (AG), one of the developers of EMS-95, will take the lead of the emissions modeling task and work closely with CARB to assure consistency and compatibility with the CARB's emissions development efforts. AG is currently under contract with CARB to develop on-road mobile source VMT estimates for the CCOS domain. AG will provide the EMS-95 system and episodic databases to the District and train the staff on its use.

The RAMS prognostic meteorological model has been selected for the modeling system because of its demonstrated successful application in the Bay Area in the past, its inclusion of all the technical features necessary for simulating the complex Bay Area meteorology, and its familiarity to the District. The CARB is currently using the MM5 meteorological model for their CCOS modeling, and is seriously considering the use of RAMS as well. However, we believe that the RAMS model may be superior for simulating the Bay Area meteorology because of its formulation and its flexibility, such as supporting any grid meshing ratio. Subcontractor Dr. Craig Tremback of ATMET is one of the developers of RAMS and will lead this task. ATMET will configure RAMS for optimal high speed performance on the project's computer cluster and train the District on its use.

The CAMx photochemical grid model was selected for the modeling system as it is publicly available, contains all of the technical options needed to simulate ozone in the Bay Area, and contains some superior capabilities to the other state-of-science models. In particular, the CAMx contains several "probing tools", including the decoupled direct method (DDM) of sensitivity evaluation, ozone source apportionment technology (OSAT), and Process Analysis, all of which will increase the likelihood of obtaining a photochemical base case simulation that fully achieves the model performance objectives. The CAMx modeling will be led by its developers, ENVIRON, who have set up CAMx for numerous ozone SIP modeling databases including Los Angeles. The CAMx modeling domain will be based upon current MM5/SAQM/CMAQ/CAMx modeling being undertaken by the CARB. However, through consultation with the District and the MAC, the specific domain configuration for the SFBA will include a nested grid arrangement that provides a balance between adequately treating the Bay Area ozone problem and addressing regional transport to and from the SFBA. ENVIRON will set up and evaluate CAMx for two CCOS episodes in 2000 and a third historical episode in July 1999, operate the model for several future year (2006) emission sensitivity tests and the final attainment demonstration, and transfer and train the District on its use.

Project-Specific Computer and Web/FTP Sites

The District has very specific computer performance goals for operation of simultaneous meteorological and air quality model simulations. As part of the current study, the project team will work with District staff to identify an appropriate Unix/Linux multi-node cluster system, purchase and test the system, install all models and supporting databases, and ultimately deliver the system to District offices near the completion of the project. This system can also be enhanced and expanded as new technology becomes available.

As a key part of the work being carried out by ENVIRON with the BAAQMD staff to develop a 2004 SIP revision, we have engaged the members of the MAC to assist in the technical review and guidance of the project. It was agreed that we would keep all project participants advised of the project through a web site maintained by ENVIRON, and to which we would post new information for the MAC as it is developed. Because this is a project site, and intended to be a resource primarily for those that are expected to make meaningful contributions to the technical review effort, we have provided password protection.

Link: http://www.environ.org/basip2004
Username: basip2004
Password: goldengate

Links to small documents such as reports, meeting minutes, summaries, and some plots will be provided via the web site to ease dissemination. However, there will likely be a need to provide larger databases to various participants during the project. In that case, we will also establish a password-protected FTP site for this purpose.

Relationship to CCOS

As noted above, the California Air Resources Board has led an effort spanning several years and expending millions of dollars, to develop a robust and highly credible data base to be used in photochemical modeling in much of California. The CCOS project is beginning to yield results in the form of a very large data set that can be used by those engaged in photochemical modeling. Data was collected for several groups of days ("episodes") in 2000. As discussed elsewhere, a subset of those episodes will be evaluated for use in this study.

Because of the complexities of photochemistry, emissions, and meteorology, it is important that air pollution management decisions be based on the best science and most comprehensive data available. While the CCOS study has moved forward significantly in many areas, the schedule under which that project is being performed is not wholly consistent with that of the BAAQMD or EPA for the 2004 SIP submission. Thus, this study, while drawing upon the CCOS effort wherever possible, will have to pursue certain technical areas on a different schedule. This is expected to affect three key areas of this study.

First, as noted previously, the base and future year CCOS emission inventories may not be completed in a timely basis for explicit use in this Bay Area study. Second, this study is using the RAMS meteorological model for reasons stated earlier. The CCOS project is using the MM5 meteorological model, but may include RAMS at some point. If MM5 modeling results

are available in a timely fashion, they will be compared to those of the RAMS model and in turn to the meteorological measurements against which the model(s) performance will be judged. Lastly, the CCOS study has not yet determined which photochemical model may be used in that study. We will be using CAMx in this study, again, for the reasons stated above.

Not withstanding the potential for different approaches in some aspects of these studies, the Bay Area study participants and the CCOS sponsors have agreed to share all technical information as both studies proceed, thus minimizing any differences in results that cannot be accounted for in an objective fashion.

Considerations for Bay Area SIP and Pollutant Transport

Pollution does not respect political boundaries. There is documented air mass flow from the Bay Area into inland areas of the State, and vice-versa. The Federal Clean Air Act recognizes such transport and addresses the manner in which up and down-wind areas are interconnected in the regulatory process. This study will provide information that should assist in the regulatory assessment. In addition to air mass and pollutant flow, there are also mobile source emissions that originate within one area but continue as vehicles move to another area. For example, automobiles registered in the San Francisco Bay Area clearly transit outside of that area. The same is true of vehicles registered elsewhere.

Both of these phenomenon will be addressed in this study. Pollutant mass transport will be explicitly addressed because the modeling domain used in this study will extend well beyond the SFBA, thus accounting for such air mass movement within the modeling system. The movement of vehicles will be addressed through the use of complex transportation model output results being used in estimating mobile source emissions. Such transportation models are used by Metropolitan Transportation Commission (MTC) in the San Francisco area, and Sacramento Association of Governments (SACOG) in the Sacramento area.

SCHEDULE AND DELIVERABLES

Figure 1-1 displays a time-line for the completion of each of the nine project tasks. The figure also shows major deliverables, to include task reports at the completion of each (except Task 2, Computer Acquisition), draft and final project reports, and the delivery of the computer system. All technical modeling development work is expected to be completed by July 1, 2003, while all future year analyses, SIP technical documentation, and transfer/training to be completed in September 2003. Table 1-1 provides the more general BAAQMD timeline for development, public review, and submission of the 2004 SIP.

The current CARB schedule of deliverables for the SFBA 2004 SIP revision effort has been delayed substantially from that estimated by CARB at the beginning of this effort, but appears to be "workable" (i.e., still allowing for a timely SIP revision to U.S. EPA in spring 2004) for the ENVIRON/District team if no further slippage is experienced.

Both the Bay Area and the CARB have identified the July/Aug 2000 CCOS episode as the first priority. The majority of emissions and meteorological data is expected to be made available

for that episode during the mid-November to early December 2002 period. Because the CARB has the June 2000 CCOS episode as their third priority, we do not expect to have final emissions or meteorological data for that episode until late December at the earliest. Nevertheless, CARB has agreed to provide the team with preliminary data such that modeling of that episode can begin and any data gaps that may exist can be determined and addressed by the team. Lastly, the CARB has strong interest in developing the July 1999 episode as part of their ozone SIP planning efforts, but it is of the lowest priority at this time. We are unable to determine what data CARB might be able to supply for that episode, nor when it would be made available. However, CARB has stated that they will be able to provide schedule and content description information for that episode shortly. The team will continue to hold discussions with CARB staff on this issue.

Although the ENVIRON/District team will have to do more work as part of this effort than originally expected and reflected in the current project "scope of work," the team is optimistic that the presently planned efforts of both groups will be sufficient to complete the Bay Area SIP in a timely fashion. Several unknowns (especially the CARB effort toward the July 1999 episode) may still significantly alter the schedule or scope of work that needs to be carried out by the ENVIRON/District team. This uncertainty, as it relates to the emissions and meteorological aspects of this project, is indicated in Figure 1-1 with dashed extensions to Tasks 3 and 4. This could have implications on the remaining tasks as well.

STUDY PARTICIPANTS

Sponsors

Bay Area Air Quality Management District
939 Ellis Street
San Francisco, CA 94109
Main: 415/771-6000
www.baaqmd.gov

Peter Hess	415/749-4971	phess@baaqmd.gov
Tom Perardi	415/749-4995	tperardi@baaqmd.gov
Rob DeMandel	415/749-4654	rdemandel@baaqmd.gov
Tom Umeda	415/749-4663	tumeda@baaqmd.gov
Phil Martien	415/749-4660	ptmartien@baaqmd.gov
Steve Soong	415/749-4657	ssoong@baaqmd.gov

Contractor Team

ENVIRON International Corporation
101 Rowland Way, Suite 220
Novato, CA 94945-5010
Main: 415/899-0700
Fax: 415/899-0707
www.environcorp.com
www.camx.com

David Souten	Principal-in-Charge	415/899-0711	dsouten@environcorp.com
Chris Emery	Project Manager	415/899-0740	cemery@environcorp.com
Ralph Morris	Advisor, Modeling and Emissions	415/899-0708	rmorris@environcorp.com
Greg Yarwood	Advisor, Chemistry and Probing Tools	415/899-0704	gyarwood@environcorp.com
Gary Wilson	Computer Acquisition, web/FTP administration	415/899-0719	gwilson@environcorp.com
Mary Mahoney	Administrative Assistant	415/899-0738	mmahoney@environcorp.com

ATMET, LLC
P.O. Box 19195
Boulder, CO 80308-9195
6610 Gunpark Drive, Suite 100
Boulder, CO 80301
Main: 303/530-5200
Fax: 303/530-5201
www.atmet.com

Craig Tremback	Meteorological Modeling	ext. 205	tremback@atmet.com

Alpine Geophysics, LLC
7107 N. Flora St. (Contracting/accounting)
Arvada, CO 80004
work: 303.425.5495
fax: 303.425.5495
7691 Alpine Rd. (Jim Wilkinson)
La Honda, CA 94020
www.alpinegeophysics.com

Jim Wilkinson	Emissions Modeling	650/854-4705	jwilkins@themis.ce.gatech.edu

Independent Consultant
San Jose State University
Department of Meteorology
1 Washington Square
San Jose, CA 95192
Main: 408/924-1000
www.sjsu.edu

Robert Bornstein	Advisor, Meteorology	408/924-5205	pblmodel@hotmail.com

The contracting team includes an Internal Technical Review Team of noted experts in emissions, meteorology, chemistry, and the development/use of model probing tools. Members will serve as an integral part of the study by reviewing the modeling as it is performed, serving as a continuous technical resource, and providing unique expertise to the contracting team. Besides the Principal-in-Charge and Project Manager, the Internal Technical Review Team includes Ralph Morris and Greg Yarwood of ENVIRON, and Robert Bornstein of San Jose State University.

Modeling Advisory Committee (MAC)

The District has formed a technical modeling advisory committee for this study. Both the EPA and CARB modeling guidance requires the formation of a "Technical Working Group". In any study of this type that leads to a SIP development effort, it is important that the technical underpinnings of that SIP be fully examined as they are developed. In this manner, to the extent possible, technical issues can be put aside in the SIP development effort, the focus can be on the many other aspects of this process, and the public can be better assured that the technical community has been rigorous in their review of the work.

The table below lists the current members of the project MAC with their contact information. The MAC includes representatives from other governmental agencies, environmental groups, and industrial stakeholders.
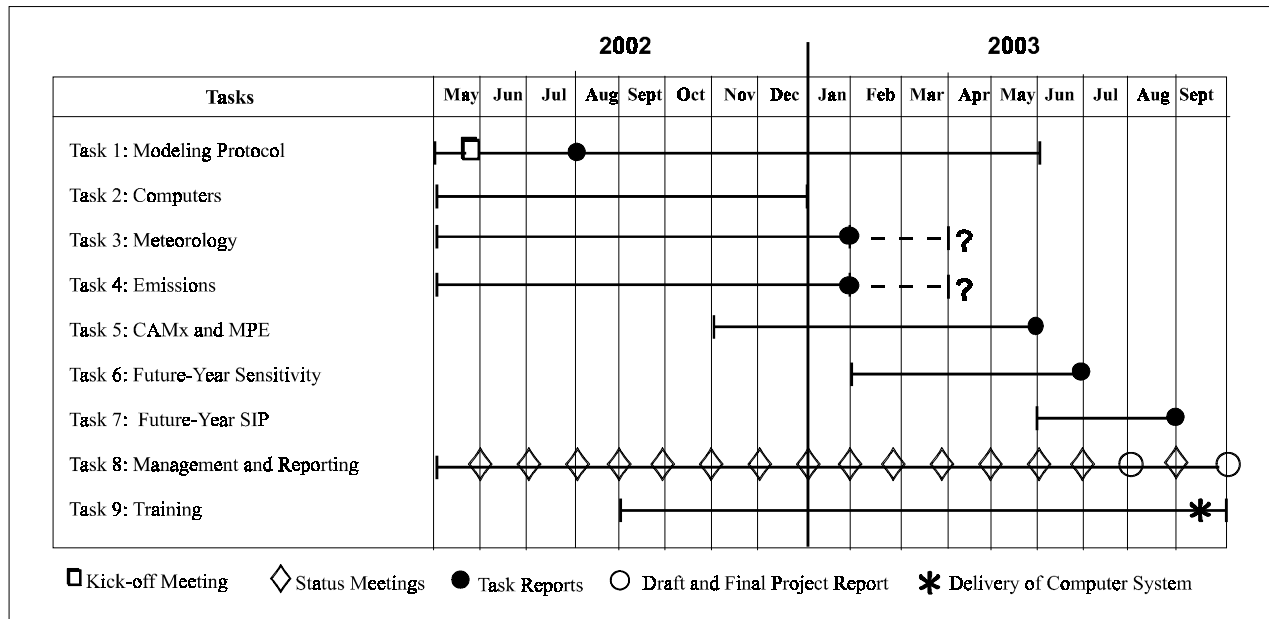
Figure 1-1. Project schedule showing dates for major deliverables.

Table 1-1. BAAQMD tentative schedule for the SFBA 2004 ozone attainment plan.

November 2002	Complete project protocol
December 2002	Begin photochemical model trial runs with available data
March 2003	Complete preparation of inputs (emissions, meteorology, etc) for base year modeling
March 2003	Preliminary future year inventory and photochemical modeling
March 2003	Regional agencies hold public meeting on development of 2004 Plan and CEQA document
June 2003	Finalize future year inventories and photochemical modeling. Derive emission reduction target.
June 2003	Regional agencies hold public workshop on technical analysis and emission reduction target. Regional agencies report to Boards on modeling results and ozone monitoring record
July 2003	Regional agencies hold public workshop on control strategy for 2004 Plan
September 2003	Release Draft Plan and Draft EIR for public review period
October 2003	Regional agencies hold public workshop on Draft Plan and Draft EIR
January 2004	Publish notice of public hearing on 2004 Plan Release Proposed Final 2004 Plan and EIR
February 2004	Hold public hearing on 2004 Plan and EIR; certify EIR
March 2004	Regional agencies adopt 2004 Plan
April 2004	ARB approves 2004 Plan and submits to EPA

Modeling Advisory Committee (MAC) Participants

Sam Altshuler Supervisor, Air Quality, Acoustics, & Industrial Hygiene PG&E Technical and Ecological Services SLA2@pge.com Phone: (925) 866-5879 Fax: (925) 866-5681	**Rob Harley** Advisory Council, BAAQMD Professor, University of California at Berkeley harley@ce.berkeley.edu Phone: (510) 643-9168 Fax: (510) 642-7483
Carol Bohnenkamp Regional Modeling Contact U.S. EPA, Region IX bohnenkamp.carol@epamail.epa.gov Phone: (415) 947-4130 Fax: (415) 947-3579	**Eugene Leong** Executive Director Association of Bay Area Governments (ABAG) eugenel@abag.ca.gov Phone: (510) 464-7910 Fax: (510) 464-7970
Harold Brazil Air Quality Associate Planner/Analyst Metropolitan Transportation Commission (MTC) hbrazil@mtc.ca.gov Phone: (510) 464-7747 Fax: (510) 464-7848	**Julia May** Clean Air Program Director Communities for a Better Environment (CBE) jmay@cbecal.org Phone: (510) 302-0430 x300 Fax: (510) 302-0437
Henry Clark West County Toxics Coalition henryc11@prodigy.net Phone: (510) 232-3427 Fax: (510) 232-4111	**Bob Nunes** Air Quality Planner II Monterey Bay Unified Air Pollution Control District (MBUAPCD) bnunes@mbuapcd.org Phone: (831) 647-9411 Fax: (831) 647-8501
John DaMassa Chief, Modeling and Meteorology Branch California Air Resources Board (CARB) jdamassa@arb.ca.gov Phone: (916) 322-6048 Fax: (916) 327-8524	**Evan Shipp** Supervising Air Quality Meteorologist San Joaquin Valley Air Pollution Control District (SJVAPCD) evan.shipp@valleyair.org Phone: (559) 230-5800 Fax: (559) 230-6040
John Holtzclaw Sierra Club john.holtzclaw@sierraclub.org Phone: (415) 977-5534 Fax:	**Brigette Tollstrup** Division Manager Sacramento Metropolitan Air Quality Management District (SMAQMD) btollstrup@airquality.org Phone: (916) 874-4832 Fax: (916) 874-4899
Larry Greene Air Pollution Control Officer Yolo-Solono Air Quality Management District (YSAQMD) lgreene@ysaqmd.org Phone: (530) 757-3656 Fax: (530) 757-3670	**Steve Ziman** Senior Staff Scientist Western States Petroleum Association ChevronTexaco Energy Research and Technology sdzi@chevrontexaco.com Phone: (510) 242-1530 Fax: (510) 242-5577
Ken Kloc Professor, Golden Gate University kkloc@ggu.edu Phone: (415) 369-5352 Fax: (415) 896-2450	

2. EPISODE SELECTION

The BAAQMD ozone SIP will include an evaluation of the effectiveness of emission control measures by simulating their effects on ambient ozone air quality. A three-dimensional air quality model, the Comprehensive Air Quality Model with extensions (CAMx), will be used for these simulations. Air quality models require time- and space-varying inputs of emission and meteorological fields. These fields significantly influence the results of the simulations and are the most relevant to the SIP analysis.

The general modeling approach for evaluating control measures is to simulate one or more historic episodes (periods that violated the air quality standard) using inputs that best approximate the physical conditions that prevailed during each episode. This simulation defines a base-year reference or "base" case. If the performance of the historic base case simulation is acceptable, then a simulation is performed using emissions that incorporate best-estimate growth projections and adopted control programs into a future year (usually the attainment year, or 2006 in this case). This "future base" case is then analyzed to indicate if any additional controls are necessary to ensure attainment of the ozone standard. If necessary, then simulations are performed using emissions that introduce proposed new emission control measures ("future control" case). The differences between the "future base" and "future control" simulations represent the air quality impacts of the proposed new emission control measures.

Episodes used for this analysis need to be selected carefully so that the analysis has the maximum credibility and generality. The criteria for episode selection are:

- The episode must have had an ozone value that exceeded the 1-hour federal ambient air quality standard. The standard for ozone is 124 ppb averaged over one hour. Ozone observations above this standard may influence the calculation of the "ozone design value", which is the regulatory measure of ozone levels.

- The episode must be representative of a class of episodes that occur frequently so that the simulation will presumably have greater generality to the analysis of predicted changes in the design value. Incorporating multiple episodes into the analysis will further broaden its generality. EPA guidance requires the examination of three or more episodes.

- The episode must have sufficient observations to determine the physical conditions that contribute to the ozone exceedances. Furthermore, the observations must provide data that satisfy model input needs and that can be used to evaluate model performance.

BAAQMD staff have investigated the categorization of ozone exceedances in the Bay Area for the period 1995 through September 2002 in order to find representative exceedance days to be used for SIP modeling. Two main categories of exceedance patterns were found: (1) when high ozone values occurred at isolated individual sites; and (2) when high values occurred at several sites and in many regions. This is discussed further in the following section.

EPISODE REPRESENTATIVENESS

Criteria For Modeling Episode Representativeness

A key question that emerges from modeling a small set of ozone exceedance days is whether the physio-chemical dynamics are adequately similar to the broader set of episodes that the conclusions derived from modeling would apply in general. That is, can the modeled days be considered *representative* of other exceedance days. One important aspect of this question is whether episode days fall into more than one clear-cut category. Do the dynamics vary sufficiently so that the modeled results of a day in one category cannot be extrapolated to days in the other category?

We are not aware of a single omnibus approach to address these issues. This analysis looks at various approaches, including simple tabulations; frequencies of episode days by day of week, month, and area; an analysis of trends by site; and a cluster analysis. The ultimate choice of modeling days necessarily involves judgment based on our experience with previous modeling, and on our conceptual understanding of ozone dynamics in the Bay Area.

Definitions

It is often a good idea to attempt to pin down the definitions of key words in an analysis. It can lead to a clarification of thought. The word "representative" is frequently used, but is a difficult one to define in this context. Let's try to answer the following: representative *of what*? Are we talking about representative *days* or representative *episodes*?

When we say "representative," it seems reasonable to assume we mean "representative *of Bay Area days exceeding to the 1-hour standard*." Because the focus of the Bay Area's ozone problem is in the east, Livermore specifically, we may want to restrict this to "representative *of eastern Bay Area days exceeding the 1-hour standard*," excluding the few days where peak ozone occurred elsewhere. We may also determine that there is more than one category of days we wish to consider, e.g., two sets of days with distinctly different dynamics. Then, we could define a day as "representative *of days exceeding the 1-hour standard that fall into category x*."

Modeling is done on episodes rather than individual days. However, episodes are more complicated because the number of days can vary, making episode to episode comparisons difficult. Also, the standard is written in terms of days, not episodes. Thus, it may be preferable to consider representative *days*. When an episode is being considered for modeling, then, we look at the days it contains to see which, if any, are "representative."

Thus, this analysis will focus on days that exceeded the 124 ppb national 1-hour ozone standard somewhere in the District. To obtain a large enough sample, we use data back to 1995. Going back further would increase the sample size, but it is unclear whether the emissions patterns back then would be sufficiently similar to current patterns.

Tables 2-1 through 2-3 present features of the 36 days from 1995 through September 17, 2002 that exceeded the national 1-hour ozone standard of 124 ppb. Table 2-2 shows numbers of exceedances by site. One feature that stands out is Livermore, which had exceedances on 27 of the 36 days, and had the highest ozone on 20 of those days. Concord was a distant second with exceedances on 10 days and the highest ozone on 6. These sites, along with Bethel Island and Fairfield, account for all but seven of the highest ozone values.

Table 2-3 shows that at least one eastern site had an exceedance on 33 out of the 36 episode days. Santa Clara County had 11 such days. The other regions – North Counties, Central Bay sites and South Central Bay sites -- have few exceedances by comparison. Table 2-3 also shows that most exceedance days (26 out of 36) occurred only in one region. Nevertheless, there were 10 days where exceedances occurred in more than one region, that is, more than once a year on average. Finally, the table shows that both 1-day and multi-day episodes are common: 16 of the 1-hour exceedances are 1-day events, the rest are 2-day or 3-day events.

Exceedances By Day of Week

Recent history suggests that ozone exceedances occur more frequently on weekend days than weekday days – a so-called "weekend effect." Figure 2-1 shows a histogram of exceedance frequencies by day of week. The average frequency is $36/7 = 5.1$, so that the weekend exceedances are somewhat more frequent than average. The differences are not statistically significant, however, and in fact, Monday has been the most common exceedance day.

The weekend effect appears strongest for inner-bay sites. Blanchard and Fairley (2001) showed that sites ringing the bay had statistically significant weekend effects whereas sites further east did not. Note that the three days when Fremont and San Jose recorded the District's maximum were weekend days and eight of the nine exceedance days for San Jose, Fremont and Alum Rock occurred on weekends. In contrast, many of Livermore's and Concord's exceedances occurred during weekdays (16 in Livermore), so weekday exceedances cannot be ignored. Eleven of Livermore's exceedances occurred on the weekends, an event that would occur only about 12% of the time by chance if the probability of a weekend exceedance was equal to that for weekdays, based on a binomial distribution. Thus, Livermore may indeed have a greater probability of an exceedance on a weekend day than a weekday.

Another way to compare the importance of weekends vs. weekdays is to compare design values. For this, seven years were used, 1996-2002, so that there would be five years of weekdays and two years of weekend days. Design values can be estimated from the 6th and 3rd largest values respectively. Typically, the design value is the 4th highest value from 3 years of data. The 7 years of data include 5 years of weekdays and 2 years of weekends. Thus, it seems reasonable to estimate the weekday design value by the 6th highest weekday value, and the weekend design value by the 3rd highest weekend value. For Livermore, these were 138 ppb for the weekdays vs. 137 ppb for the weekends.

Table 2-1. Ozone (ppb) by site on District days exceeding the national 124 ppb 1-hour ozone standard, 1995 through September 2002. Exceedances in bold; District max underlined. Successive episodes shown by alternating shading/no shading.

m/d/yr	dow	Eastern						North				Central Bay			South Central Bay					Santa Clara Valley					BA Max	SJV Max	SAC Max
		bi	ff	pt	cc	li	lv	np	va	st	sr	oa	sf	ri	mv	rc	sl	fr	ha	lg	gi	sm	ar	sj			
6/23/95	Fr	86	95	96	90	**130**		81		53	60	36	26	49		68	77	108	88	89	118	113	**127**	100	**130**	142	116
6/24/95	Sa	76	109	87	98	**142**		74		67	84	84	74	81		116	**150**	**153**	**145**	**130**	107	115	**127**	121	**153**	133	106
6/25/95	Su	98	**129**	121	**128**	120		**130**		84	88	75	71	78		114	**131**	117	119	97	92	91	113	114	**131**	132	125
7/14/95	Fr	79	95	91	86	106		80	82	61	57	33	53	52	97	88	67	92	102	101	**130**	**128**	113	108	**130**	128	111
7/15/95	Sa	84	87	93	99	98		90		88	82	70	88	87	116	**140**	**144**	**149**	**138**	**128**	108	107	**145**	**134**	**149**	140	117
7/27/95	Th	124	113	119	**152**	**155**		105	91	97	59	26	40	47	79	58	87	107	88	**141**	96	102	106	102	**155**	156	131
7/31/95	Mo	104	104	104	121	**138**		78	71	73	54	30	40	62	90	69	82	87	94	**135**	86	93	113	98	**138**	149	154
8/14/95	Mo	107	98	113	**147**	134		95	99	73	68	34	42	57	66	60	87	100	88	85	81	88	105	98	**147**	139	111
8/19/95	Sa	74	88	86	92	**147**		90	100	76	73	75	58	79	84	103	100	99	101	107	90	86	102	97	**147**	129	123
8/20/95	Su	73	63	71	75	**130**		53	43	52	30	27	37	36	60	56	59	75	64	101	57	69	94	90	**130**	140	116
9/795	Th	**128**	95	124	92	78		92	61	66	47	21	42	53	52	48	69	68	77	63	90	91	80	59	**128**	118	114
6/3/96	Mo		81	75	87	**128**		73	83	53	51	21	20	27	42	32	38	71		72	91	82	73	59	**128**	126	113
6/30/96	Su	79	100	92	115	**131**		90	100	66	80	85	51	73	94	69	107	89		113	90	114	88	90	**131**	137	114
7/1/96	Mo	**137**	113	117	**127**	**133**		83	93	83	56	30	47	44	80	49	79	90		**129**	95	94	102	88	**137**	143	126
7/21/96	Su	86	74	69	85	**126**		62	54	68	53		27	36	69	34	44	74		91	78	83	86	81	**126**	115	126
7/28/96	Su	77	89	91	95	**129**		86	69	61	46	27	25	33	68	41	55	75		92	84	103	74	72	**129**	121	93
8/8/96	Th	90	67	87	99	**133**		76	60	48	37	22	30	33	47		30	53		57	104	99	57	61	**133**	150	110
8/9/96	Fr	113	101	94	101	**138**		78	81	60	55	36	31	42	71	46	71	76		96	98	109	62	88	**138**	144	150
8/10/96	Sa		76		97	**137**		69		45		21	23	31	50	32		50		77	92		51		**137**	148	113
7/18/98	Sa	97	102	95	115	**146**		91	79	63	69	43	26	58	97	54	73	106	99	**133**	**132**	**135**	**129**	**147**	**147**	158	104
8/3/98	Mo	91	98	75	84	124		72	63	67	73	43	29	56	95	66	90	96	94		**135**	**142**	98	109	**142**	143	151
8/4/98	Tu	101	121	97	119	**134**		**125**	119	67	74	34	29	47	85	43	101	115	104		118	**144**	120	110	**144**	153	148
8/12/98	We	123	106	95	**147**	139		101	106	68	63	20	19	49	81	42	77	89	92	92	97	112	72	76	**147**	145	130
8/29/98	Sa	88	70	63	88	**131**		66	59	44	42		30	36	56	36	37	63	63	68	91	98	61	66	**131**	154	111
9/2/98	We	113	87	70	98	**139**		72	59	56	39		21	29	50	32	45	61	59	76	96	92	47	56	**139**	153	145
9/3/98	Th	120	110	96	**130**	113		101	104	66	57		23	38	81	36	111	98	116	60	63	72	86	101	**130**	119	151
9/13/98	Su	94	90	68	87	**136**		88	75	62	52		36	56	83	55	86	102	98	92	88	94	96	88	**136**	124	127
7/11/99	Su	99	117	88	**126**	**146**		105	113	76	92	76	52	67	109	69	113	**133**	123	116	104	**125**	116	103	**146**	142	137
7/12/99	Mo	112	120	98	**156**	144		115	95	76	61	27	50	47	95	47	89	98	83		101	115	107	109	**156**	132	140
8/25/99	We	**128**	**129**	95	109	94		103	98	73	81	34	26	63	100		52	70	68	117	105	110	89	109	**129**	144	160
5/22/00	Mo	115	82	107	**138**	84	82	70	65	43	48	30	21	47		29	43	55	50	46		77	61	58	**138**	139	134
6/15/00	Th	85	66	78	86	**137**	**152**	57	36	37	32	31	30	36		36	35	44		63		64	56	52	**152**	139	124
7/31/00	Mo	93	79	84	81	**126**	124	63	46	48	25	15	17	23		25	20	56	51	62		48	46	46	**126**	118	103
7/3/01	Tu	**130**	102	118	**134**		113	99	82	86	68	38	58		55	45	80	91	76	90	90	95	77	84	**134**	134	110
7/9/02	Tu	97		87	84		**135**	76	82	50	57	34	35	54	70	54	65	84	69	92	121	116	76		**135**	133	145
7/10/02	We	111	101	111	102		**160**	78	73	67	60	29	29	51	67	48	75	67	70	106	75	90	67		**160**	163	137

Sites

bi	Bethel Island
ff	Fairfield
pt	Pittsburg
cc	Concord
li	Livermore – old 1st st
lv	Livermore – Rincon
np	Napa
va	Vallejo
st	Santa Rosa
sr	San Rafael
oa	Oakland
sf	San Francisco
ri	Richmond/San Pablo
mv	Mountain View/ Sunnyvale
rc	Redwood City
sl	San Leandro
fr	Fremont
ha	Hayward
lg	Los Gatos
gi	Gilroy
sm	San Martin
ar	Alum Rock
sj	San Jose

Table 2-2. Number of 1-hour exceedances (District maximum) by site, 1995 through September 2002.

North Counties (NC)		Eastern (E)		South Central Bay (SCB)		Santa Clara County (SCC)		Central Bay (CB)	
Napa	2(0)	Bethel Isl.	4 (2)	Fremont	3(2)	Alum Rock	4(0)	Oakland	0
San Rafael	0	Concord	10(6)	Hayward	2(0)	Gilroy	3(1)	SF	0
Santa Rosa	0	Fairfield	2(1)	Mountain View	0	Los Gatos	6(0)	San Pablo /Richmond	0
Vallejo	0	Livermore	27(20)	Redwood City	1(0)	San Jose	2(1)		
		Pittsburg	0	San Leandro	3(1)	San Martin	5(2)		

Table 2-3. Number of 1-hour exceedances by year, subregion, number of areas, and episode length, 1995 through September 2002.

		Exceedances by subregion					No. of days with n regions exceeding std			Episode length		
Year	District	NC	CB	E	SCB	SCC	n=1	n=2	n=3	1 day	2 days	3 days
1995	11	1	0	9	3	6	5	4	2	4	2	1
1996	8	0	0	8	0	1	7	1	0	3	1	1
1997	0	0	0	0	0	0	0	0	0	0	0	0
1998	8	1	0	7	0	3	6	1	1	4	2	0
1999	3	0	0	3	1	1	2	0	1	1	1	0
2000	3	0	0	3	0	0	3	0	0	3	0	0
2001	1	0	0	1	0	0	1	0	0	1	0	0
2002*	2	0	0	2	0	0	2	0	0	0	1	0
All	36	2	0	33	4	11	26	6	4	16	7	2

*Through September 17, 2002
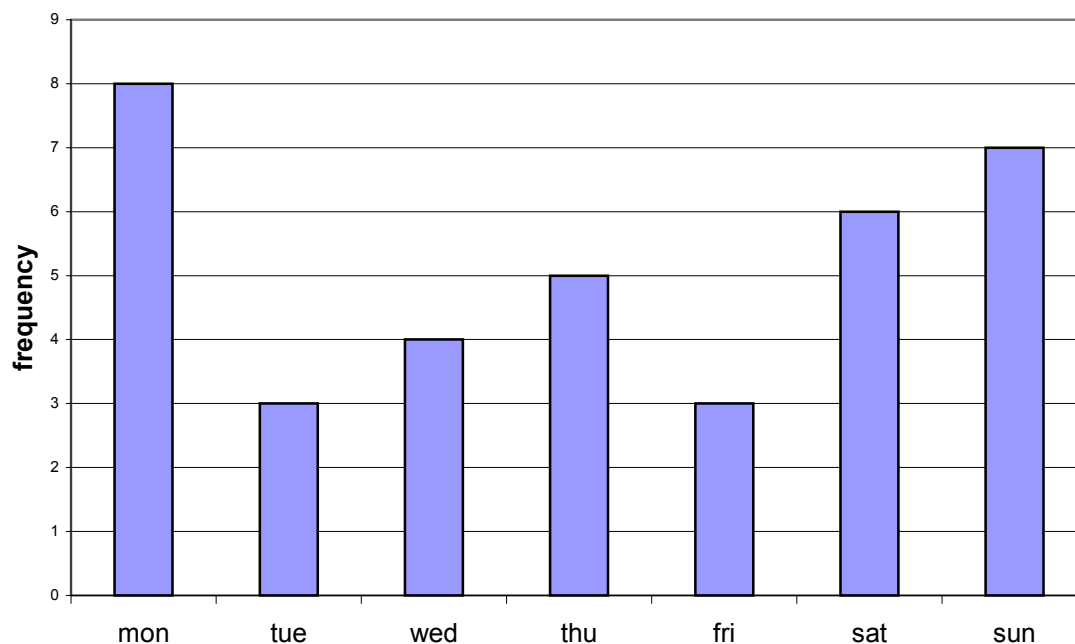
Figure 2-1. Numbers of exceedances by day of week, 1995 through September 2002.

Ninety percent confidence intervals can be constructed based on the binomial distribution.[1] They are 134 ppb to 152 ppb for weekdays, compared with 136 ppb to 146 ppb for the weekends. Thus, there is no evidence that Livermore's weekend design value is higher or lower than its weekday design value. By way of contrast, Fremont has a strong weekend effect. Its weekday and weekend design values are 98 ppb and 109 ppb respectively, with confidence intervals of (96, 106) for weekdays, and (106, 133) for weekends. This difference is statistically significant.

Month of Exceedance

In the Bay Area, ozone exceedances are most frequent in July and August (Figure 2-2). In fact, all but two of the 36 exceedances in 1995 through September 2002 occurred between June 15 and September 15.
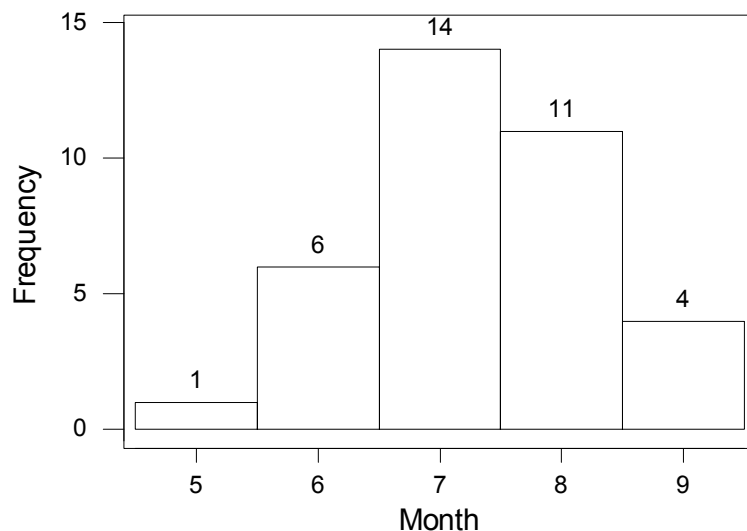
Figure 2-2. Numbers of exceedances by month, 1995 through September 2002.

This suggests that, to ensure representativeness, modeling days should be selected from this period. However, it is not clear that the dynamics are necessarily different at other times; the main reason for the fewer exceedances may simply be that there tend to be more hot days between mid-June and mid-September.

---

[1] Here a confidence interval for the upper $1/365^{th}$ percentile is shown. This is not identical to the design value however. With three years of data, this would usually be estimated as the 3rd largest value, whereas the design value is estimated as the 4th largest. Also, technically, the binomial distribution is assuming that the days are statistically independent, whereas in reality, ozone values are serially correlated. This makes the confidence intervals narrower than they would be if we could take this correlation into account.

Since ozone production depends on sunlight, the dynamics of ozone production in September and October may be different from earlier summer months. September exceedance days in recent years experienced localized ozone, with only one station exceeding the standard on each day.

Trends and Representative Days

Figure 2-3 shows numbers of exceedances by decade (1982-1991 vs. 1992-2001) for long-running BAAQMD sites. Livermore was the hot spot in both decades, but during the 1980's there were other sites that competed with it, Los Gatos in particular. What is apparent from the figure is that there has been a dramatic improvement in the south bay area, but marginal if any improvement in the east (Livermore, Concord, Bethel Island). The point is, as more previously adopted controls are implemented, one would expect more of the same, namely further improvement in the south bay, but not necessarily in the eastern part of the district. This implies that the regulatory focus should be on reducing ozone in the eastern part of the district, and hence modeling also needs to be focused on days when ozone is high in the east.

Cluster Analysis

Another way to assess representativeness is through cluster analysis. This analysis finds groups of days that are similar, based on some numeric variables. The variables used here are the 1-hour ozone maxima from the various BAAQMD sites. Two days are "similar" if the patterns of high ozone are similar. In this analysis there were 36 episode days with ozone measurements from 21 BAAQMD sites.[2]

Figure 2-4 presents the results of the analysis. Pairs of days connected by short lines are most similar. For example, 8/8/96 and 8/10/98 have lines connecting them with a distance of about 7 ppb (representing the average difference between the 1-hour values on the two days from the 21 sites). The cluster of those two days is similar to 8/29/96. The cluster containing those three days is similar to 9/2/98, and so on.

The analysis shows two days that are very different from the rest, 6/24/95 and 7/15/95, denoted as cluster 3 in the figure. These were days when the maximum ozone occurred at Fremont and there was high ozone at other sites near San Francisco Bay. We can break the remaining days into two clusters, denoted in the figure by 1 and 2. Comparing cluster 2 days with Table 2-1, we note that in every case, the maximum occurred at just one eastern site, with relatively low values everywhere else. Cluster 1 days contain all days where there was ozone at multiple sites and regions, thus representing more widespread high ozone.

Table 2-4 shows that Cluster 1 District peak ozone values were somewhat higher than Cluster 2's. From Figure 2-5, however, there was considerable overlap. Cluster 3's peaks were

---

[2] Missing values were filled in using a combination 2-way ANOVA to provide initial estimates, then assuming the values for a particular day came from a multivariate normal distribution, estimating the parameters using the E-M algorithm, then using the MVN to predict the missing values.

among the highest, but the paucity of data makes statistical inference difficult – the differences are not statistically significant.

In contrast, the means of the daily 1-hour maximum ozone values from BAAQMD sites show dramatic differences. There is no overlap between the three clusters (see Figure 2-6), with Cluster 2 having the lowest mean values, Cluster 1 in the middle, and Cluster 3 the highest. Among the six selected days shown in Table 2-4, the mean ozone values for the two from Cluster 2, 6/15/00 and 7/31/00, show the lowest mean ozone. The two 2002 days have mean ozone that is substantially below the Cluster 1 average, whereas the 1999 days are above the Cluster 1 average.

Table 2-4 also shows the daily 1-hour maximum ozone for the Sacramento and San Joaquin Valley districts. For SJV, there is no statistical difference between the clusters. For Sacramento, however, Cluster 1 days have substantially higher ozone than Cluster 2 days. Note that among the six days shown in Table 2-4, SV exceeded the 1-hour standard on all the Cluster 1 days and neither of the Cluster 2 days.

Comparisons of Meteorological Variables by Cluster

A number of meteorological variables were chosen for comparison between clusters. The basis for the choice was previous experience – variables that had been shown to be useful for predicting high ozone. These included daily maximum temperatures and midday wind speeds at several Bay Area surface meteorological monitoring sites and various RAOB measurements collected at Oakland at 4 AM and 4 PM daily. Among the temperature variables, there was some difference between Clusters 1 and 2 at Livermore, but still considerable overlap. The 850 mb temperatures exhibited no significant difference between the clusters. For San Jose maximum temperature, however, there was a clear-cut, highly statistically significant difference, with 75 percent of the Cluster 1 temperatures greater than 75 percent of the Cluster 2 temperatures.

The Cluster 1 midday winds at Travis AFB and San Martin were somewhat lighter than for Cluster 2. The 850 mb wind speeds showed no difference, nor did the 850 mb 4 AM wind direction. The 850 mb 4 PM wind direction did show clear-cut differences (Figure 2-7). On over half of the Cluster 1 days the 4 PM winds had an easterly component. In contrast, only one of the 12 Cluster 1 days (and 1 out of 2 Cluster 3 days) had an easterly component.

In summary, there are indeed differences between the clusters for some meteorological variables. Cluster 2 days tended to be cooler, especially in the south bay. Cluster 2 days had somewhat stronger winds. Almost all Cluster 2 days had westerly 850 mb winds at 4 PM, whereas over half the 850 mb winds at 4 PM for Cluster 1 had an easterly component.
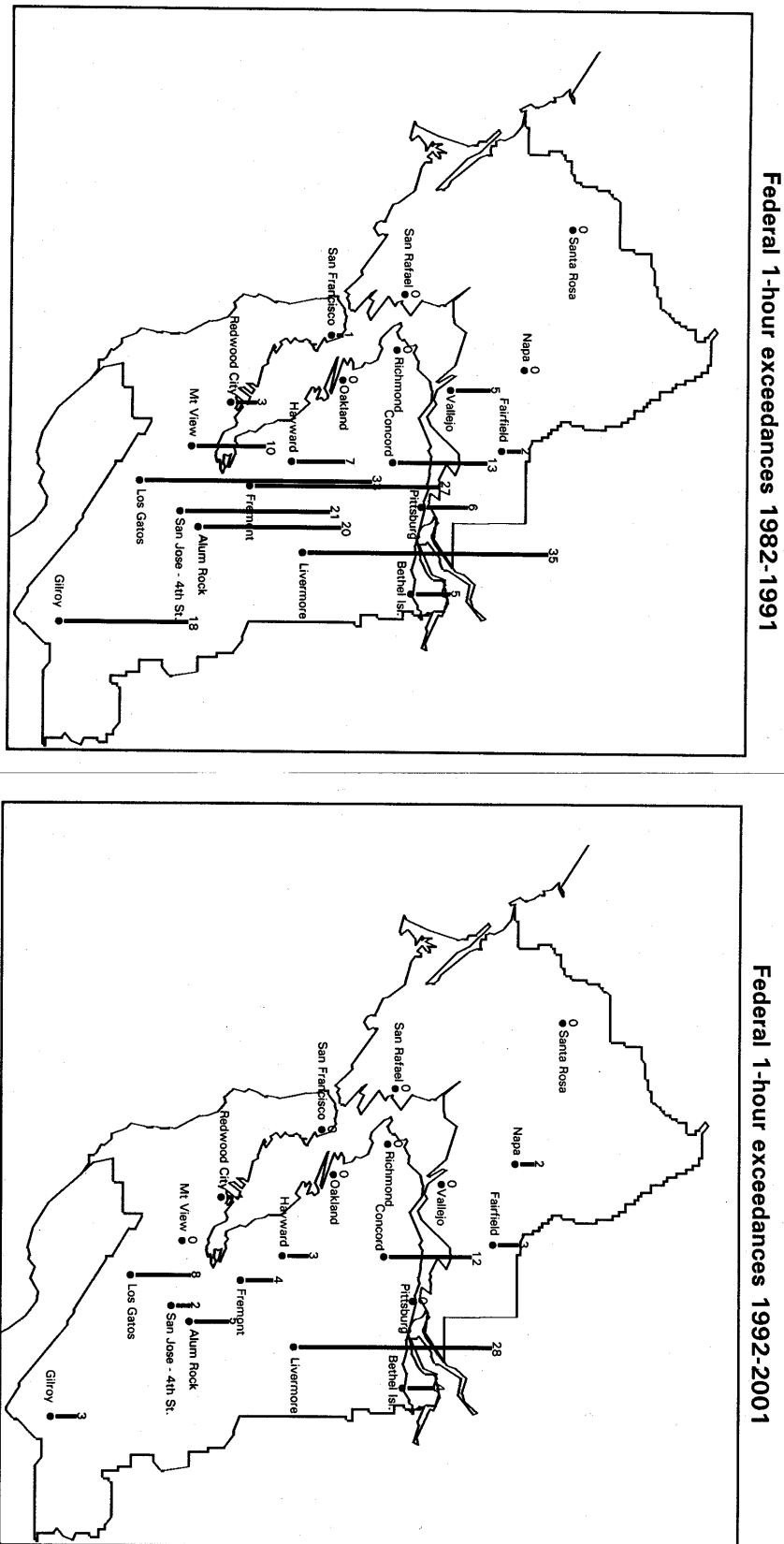
Figure 2-3. Trends in exceedances of the federal 1-hour ozone standard 1982-91 vs. 1992-2001. Vertical bars represent numbers of exceedances at long-running BAAQMD sites.
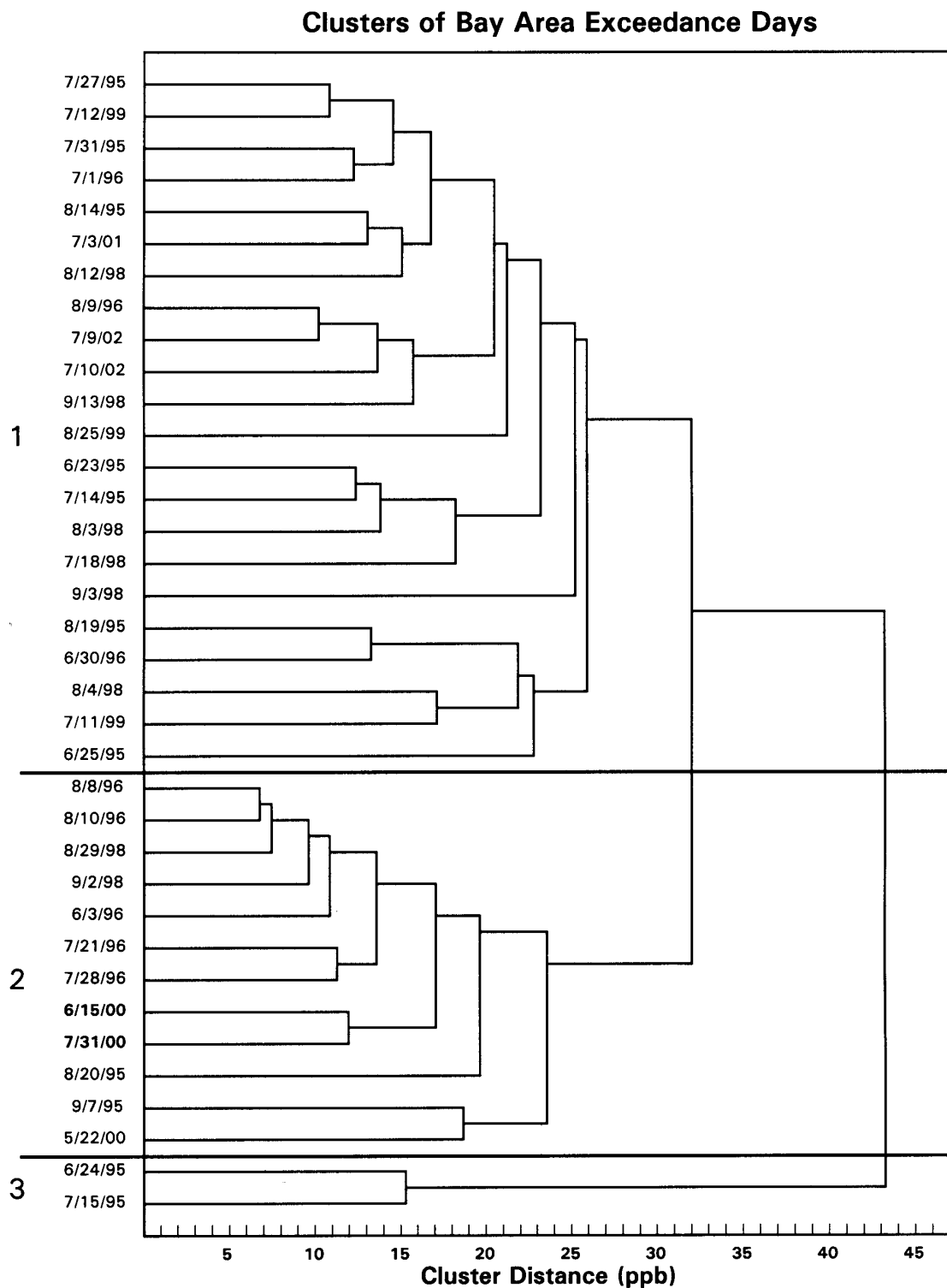
Figure 2-4. Clustering of Bay Area days exceeding the national 1-hour ozone standard, 1995 through September 2002. Thick, horizontal lines divide the three main clusters.

Table 2-4. Ozone and meteorological summary statistics: medians by cluster and values by key exceedance day.

Cluster:	1	2	3	Cluster Diffs*	7/11/99 Clust 1	7/12/99 Clust 1	6/15/00 Clust 2	7/31/00 Clust 2	7/9/02 Clust 1	7/10/02 Clust 1
n:	22	12	2							
District Max O_3 (ppb)	138	131	151	2 < 1	146	156	152	126	135	160
District Avg O_3 (ppb)	89	65	107	2 < < < 1	102.8	92.5	52.7	56.7	77.3	78.5
Livermore max T (F)	103	99	102	2 < 1	104	105	102.8	101.4	103.8	105.9
SJ max T (F)	98	92	103	2 < < < 1	95	101	90	93	101	99
850 mb 4am T (F)	77	75	71	1≈2	77	78	79	77	77	85
850 mb 4 pm T (F)	77	74	74	1≈2,3	78	81	81	79	80	85
Travis 10-4 WS (mph)	6.7	7.7	6.7	1 < 2	5.4	8.6	11.9	8.6		
S Martin 10-4 WS (mph)	8.0	9.9	7.9	1 < 2	6.5	7.9	10.2	9.5	8.4	7.8
850 mb 4 am WS (mph)	6	6	5	1≈2	14	6	18	5	5	2
850 mb 4 pm WS (mph)	7	6	3	1≈2	12	5	11	2	3	10
850 mb 4 am WD (deg)	150	235	78	1≈2	345	160	10	310	75	360
850 mb 4 pm WD (deg)	170	252	260	1 < < 2	25	245	255	280	175	300
SJV max O_3 (ppb)	142	139	137	1≈2	142	132	139	118	133	163
SAC max O_3 (ppb)	131	114	112	1 < 2	137	140	124	103	145	137

* This column shows the extent to which the clusters differed for the selected ozone and met. variables. A ≈ indicates that there was no statistically significant difference. One < sign indicates statistical significance (p < .05). Two < signs indicate significance at the .01 level. Three < signs indicates cognizance at the .001 level. Because Cluster 3 had only 2 values, it was not reasonable to test for statistical significance.
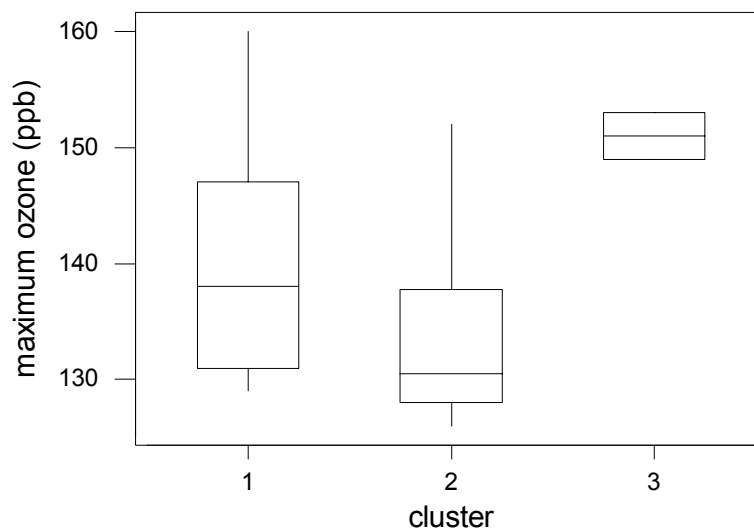
Figure 2-5. Boxplots of daily maximum ozone by cluster. Boxes (rectangles) show 25th and 75th percentiles. Horizontal lines in the boxes are medians. Vertical lines above and below the boxes indicate the range of the data unless there are outliers. Outliers are shown with asterisks (see Figure 2-6).
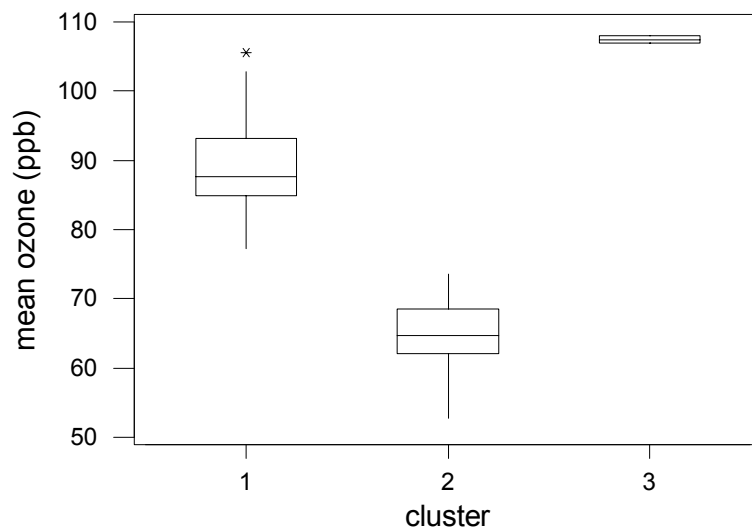
Figure 2-6. Boxplots of mean 1-hour ozone maxima for the 21 sites, by cluster, 1995 through September 2002.
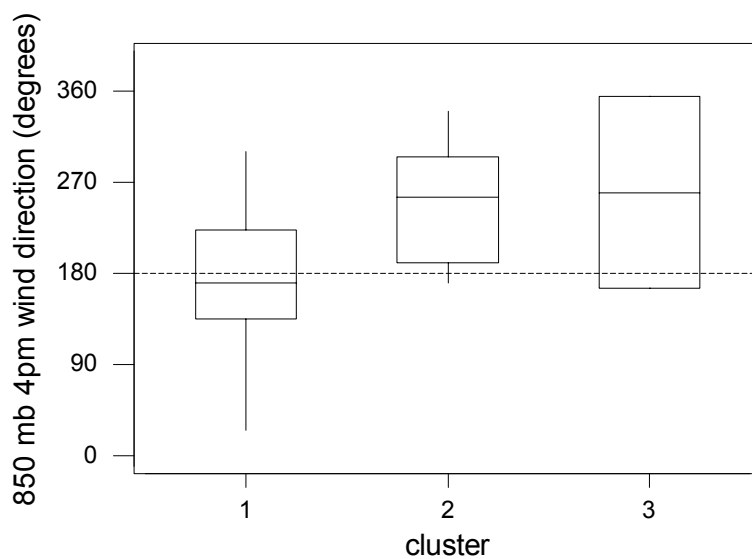
Figure 2-7. 850 mb 4 PM wind directions at Oakland by cluster, 1995 through September 2002. A line is drawn at 180° for reference.

Analysis of Meteorological Conditions

The cluster analysis in the previous section produced two distinct ozone episode clusters: Cluster 1 with ozone exceedances in several regions of the Bay Area and Cluster 2 with ozone exceedances mostly in the Concord and Livermore areas. The previous section also examined the influence of meteorological variables for 6 episode days (7/11-12/1999, 6/15/2000, 7/31/2000, and 7/9-10/2002) that are potential candidates for the Bay Area SIP modeling work. The June 15 and July 31, 2000 episodes, both Cluster 2 episode days, have been selected as the first 2 episodes to be modeled since they were part of the CCOS field study. A third episode will be selected from the Cluster 1 episodes of July 11-12, 1999 or July 9-10, 2002. The representativeness of these episodes are examined by analyzing available meteorological data.

Analysis of the weather maps showed that all the 6 episode days were characterized by a high 500 mb geopotential height and high 850 mb temperature, indicating strong downward motion. This downward motion created a strong inversion layer in the Oakland sounding. The surface weather maps also showed thermal lows or troughs over California for each of the 6 days. There was no clear distinction in the weather patterns between Cluster 1 episode days and Cluster 2 episode days.

Trajectory Analysis

Another way to characterize the meteorological conditions is to analyze back trajectories from various points in the Bay Area to identify transport routes and possible source areas of ozone precursors. Figures 2-8 and 2-9 show back trajectories from Livermore and San Martin ending at 2 PM PST on July 11, 1999, computed by HYSPLIT using the EDAS (Eta Data Assimilation System) wind fields. The surface trajectory arriving at Livermore came from the North Bay area. The surface trajectory arriving at San Martin came from the Peninsula and the Central Bay Area. Both of the trajectories passed through areas rich in ozone precursors and could be linked to the high ozone observed at Livermore and San Martin on this day.

As a matter of fact, the surface trajectories reaching Livermore at 2 PM for each of the 6 days (the other 5 days not shown) all came from the west, passing through the areas surrounding the San Francisco Bay; high ozone was observed at Livermore for all days. The surface trajectories at San Martin passed through the San Francisco Bay area on 5 of the 6 days (the other 5 days not shown). High ozone was observed at San Martin in 4 of these 5 days. July 31, 2000 was the only exception, when the observed ozone maximum was 46 ppb. On June 15, 2000, the surface trajectories at San Martin passed through the Santa Cruz Mountains and the observed ozone maximum at San Martin was 56 ppb.

Note also the vastly different trajectory paths for the three different end-point elevations (suface, 500 m, and 1000 m). This shows that a high degree of vertical shear is present during July 11. Certainly, this opens up the possibility for contributions of ozone and precursors reaching the eastern Bay Area from the Central Valley, particularly Sacramento. While the 1000 m trajectory is probably higher than the 2 PM mixing depth in Livermore, the 500 m trajectory should be near the top of boundary layer, indicating that some pollutants could be arriving in Livermore from the north and east.

Figure 2-8. Back trajectories ending at Livermore at 2 PM July 11, 1999. The red line (with triangles) is the surface trajectory at 2 m. The blue line (with square) is the trajectory at 500 m. The green line (with circle) is the trajectory at 1000 m. The time is shown in UTC. To convert to PST, subtract 8 hours from UTC.

NATIONAL OCEANIC ATMOSPHERIC ADMINISTRATION
Backward trajectories ending at 22 UTC 11 Jul 99
EDAS Meteorological Data

Job ID: 32644   Job Start: Wed Sep 25 16:51:37 GMT 2002
lat.: 37.086 lon.: -121.6 hgts: 2, 500, 1000 m AGL

Trajectory Direction: Backward   Duration: 24 hrs
Vertical Motion Calculation Method: Model Vertical Velocity
Produced using NOAA ARL Website (www.arl.noaa.gov/ready/)
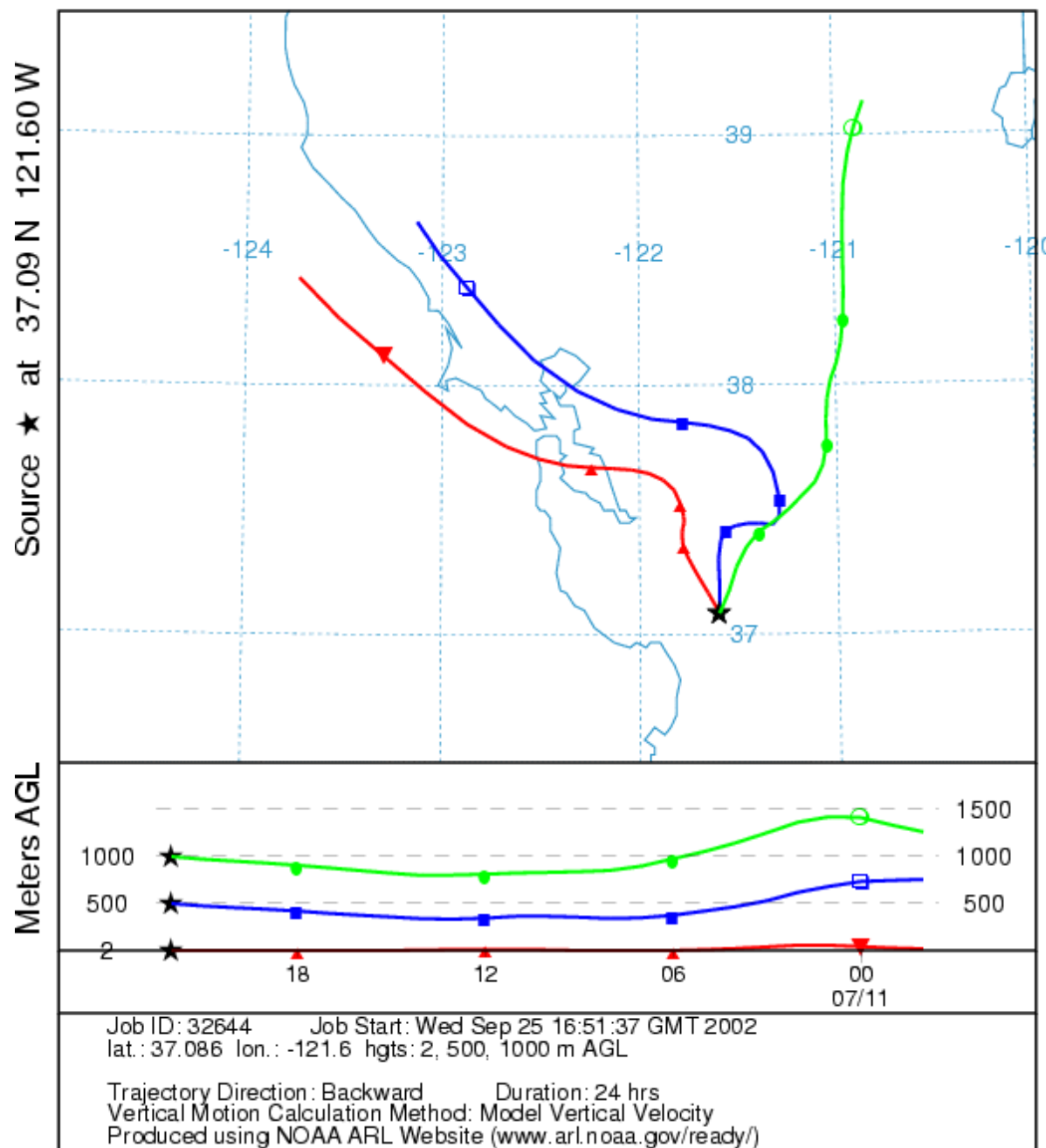
Figure 2-9. Same as Figure 2-8 except for San Martin.

We must be cautious in drawing conclusions purely based on the trajectory analysis for the following two reasons:

1. Individual trajectories vary greatly with time and height. Since the trajectories in Figures 2-8 and 2-9 are the back trajectories ending at 2 PM, when the sea breeze penetrated into the Livermore and Santa Clara Valley, the trajectories are most likely to be from the ocean. Some of the trajectories computed for the early morning hours, when the wind at Livermore was weak or from the east, originated from the Central Valley.

2. The EDAS data, having a 40 km resolution, do not resolve the intricate topography of the Bay Area. Specifically, the EDAS data may not resolve the Tri-Valley area, which is important in the assessment of transport to Livermore, as discussed in the next section.

Surface Observations

Figures 7-10 through 7-21 show surface wind observations in the Bay Area at 7 AM and 2 PM PST for the 6 episode days.

The wind patterns at 7 AM, a heavy commuting hour, are generally light and variable, but differ quite a bit among the 6 episode days. For example, the two CCOS 2000 episode days are characterized by strong winds; north-northwest winds of 8-10 MPH had already been established by this hour in the Martinez area. There were also south-southwest winds between Livermore and San Jose, forming a clear convergence in the Tri-Valley area. During the other 4 episode days, the morning winds were weak. Northwest winds existed on July 11, 1999, and on both July 9 and 10, 2002. The winds in the Martinez and Pittsburg areas funneled into the San Ramon Valley. However, on July 12, 1999, morning winds in the Martinez area were in general from the west (northerly at only one station). By 9 AM on July 12, north-northwest winds were also established in the Martinez area.

The afternoon wind patterns of the 6 days were amazingly similar in all areas except in the Santa Clara Valley. One consistent feature was the northwest to north-northwest wind near Martinez; yet another feature was the strong afternoon westerly flow at Pleasanton. The northwest wind near Martinez either persisted throughout the previous night or started in the early morning. This wind may transport ozone precursors from this area down to the Tri-Valley area over a long period of time. The wind at Pleasanton was weak in the morning; the westerly wind usually did not start until after 10 AM, and sometimes started as late as 1 PM. This sea/bay breeze could transport ozone precursors from areas surrounding the San Francisco Bay into the Livermore area via the Castro Valley-Dublin gap. Therefore, high ozone in Livermore may be due to a convergent inflow from two areas of ozone precursors. The similarities in the wind pattern on these 6 days and the occurrence of ozone at Livermore may indicate that the mechanism for ozone production in Cluster 1 and Cluster 2 days are similar. Resolving the topography and winds in the Tri-Valley area may be necessary for proper modeling of Livermore ozone formation.

The afternoon wind at San Martin was northerly for half of the 6 days and southerly for the other half. Of the 4 high ozone days at San Martin, 2 days showed north wind and 2 days showed south wind. Of the 2 low ozone days at San Martin, 1 day showed north wind and the
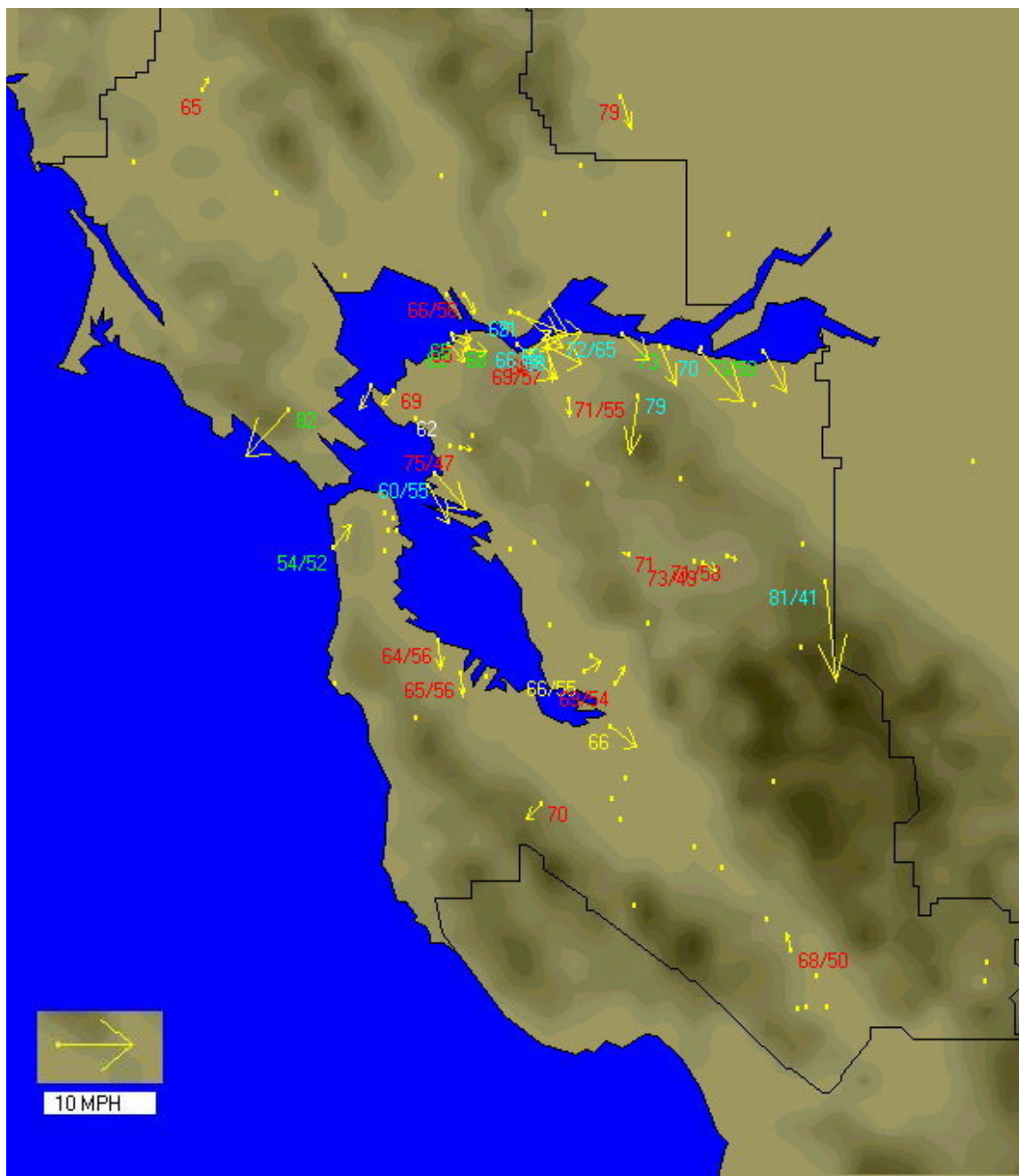
Figure 2-10. Bay area surface-wind observations at 7 AM PST, July 11, 1999. The numbers are temperature and dew point temperature.

Figure 2-11. Bay area surface-wind observations at 2 PM PST, July 11, 1999. The numbers are temperature and dew point temperature.
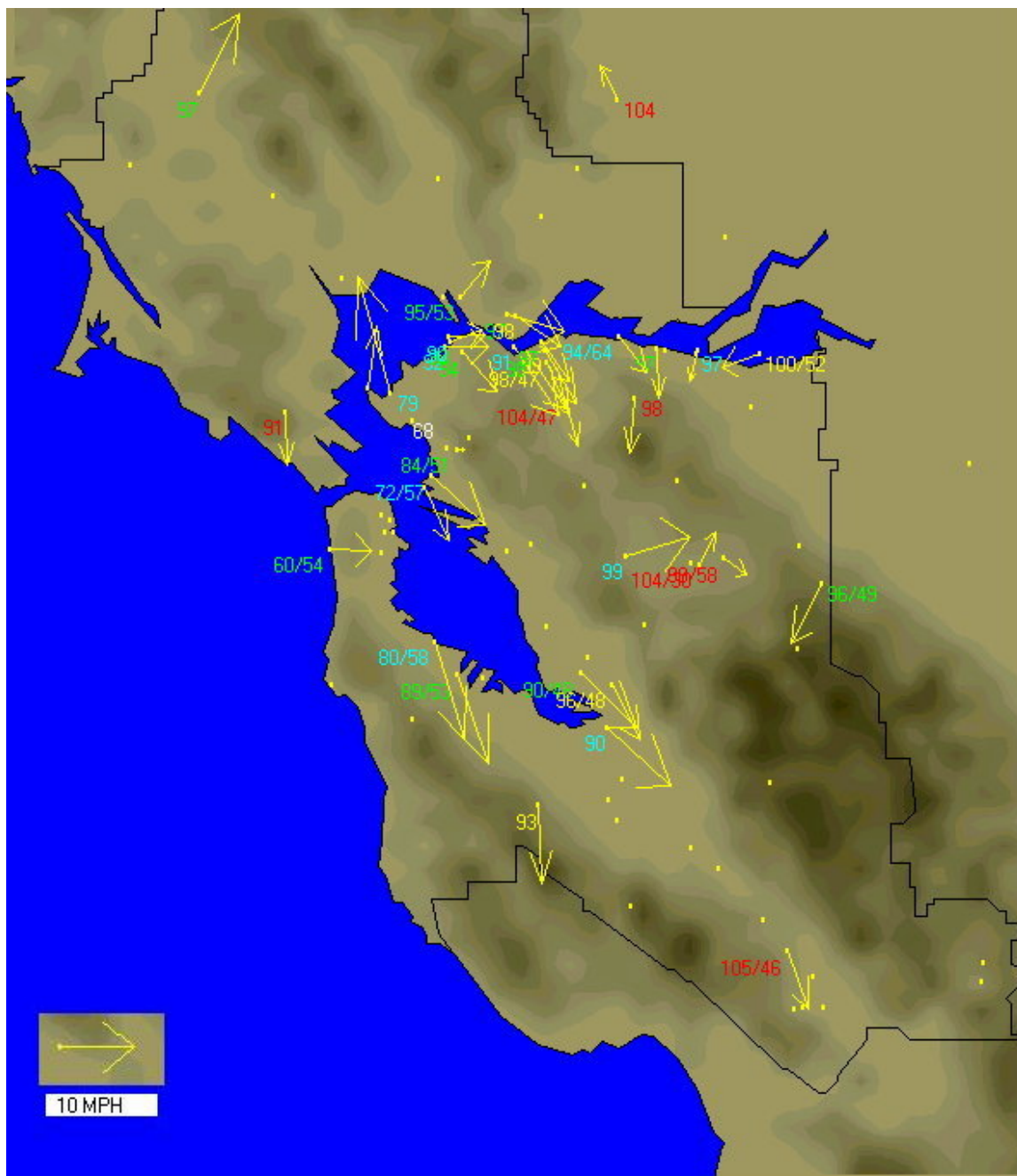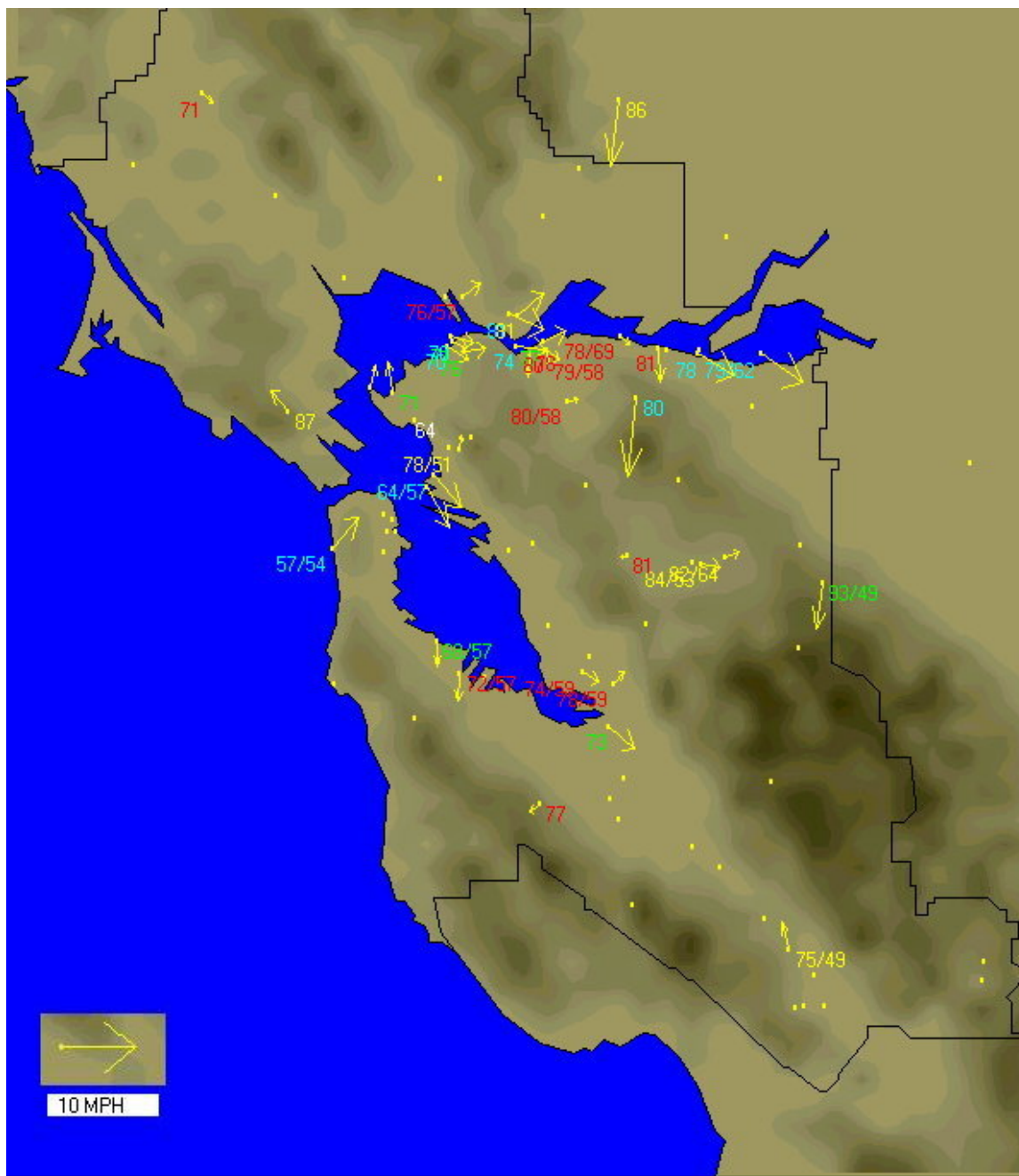
Figure 2-12. Bay area surface-wind observations at 7 AM PST, July 12, 1999. The numbers are temperature and dew point temperature.
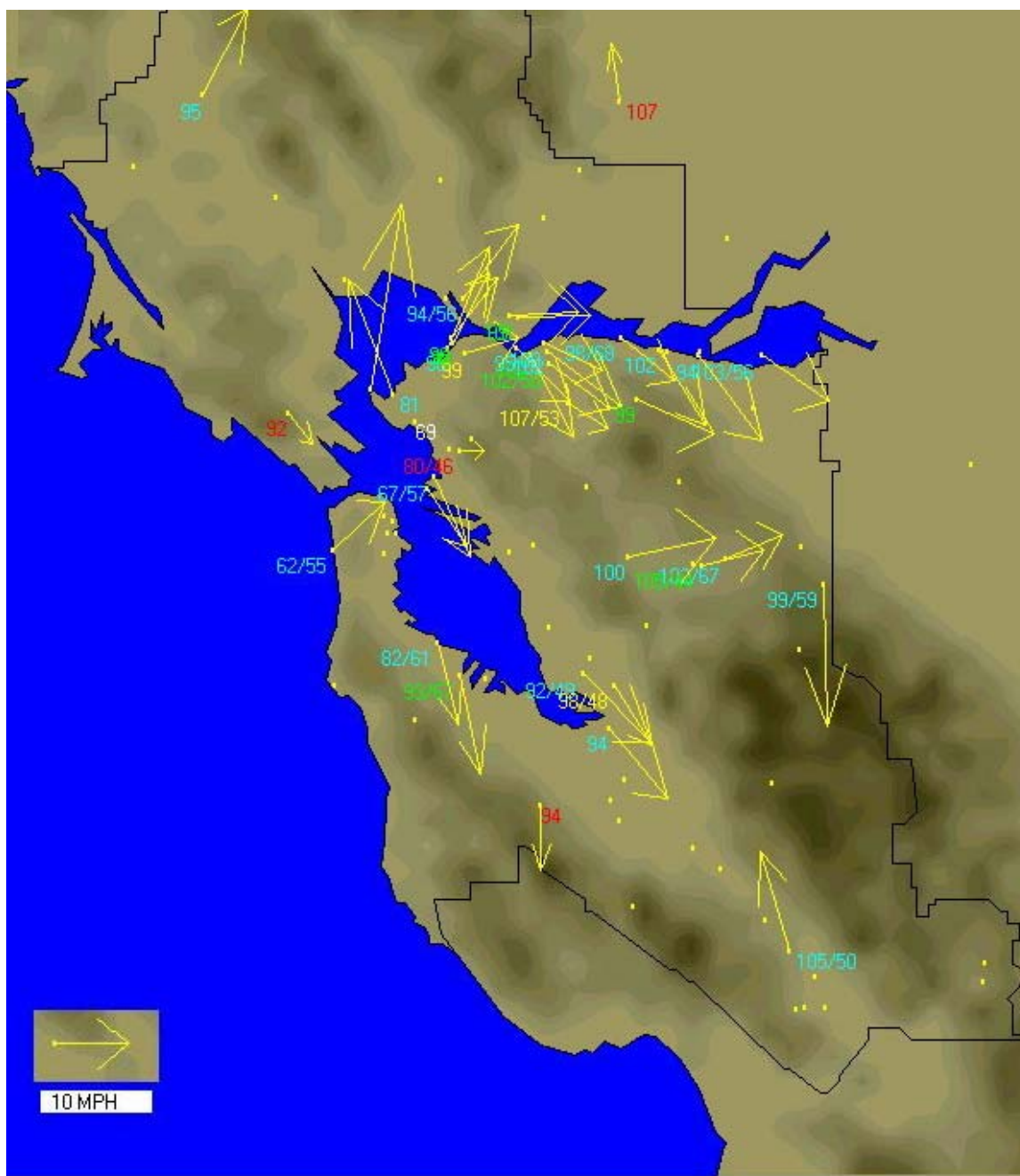
Figure 2-13. Bay area surface-wind observations at 2 PM PST, July 12, 1999. The numbers are temperature and dew point temperature.
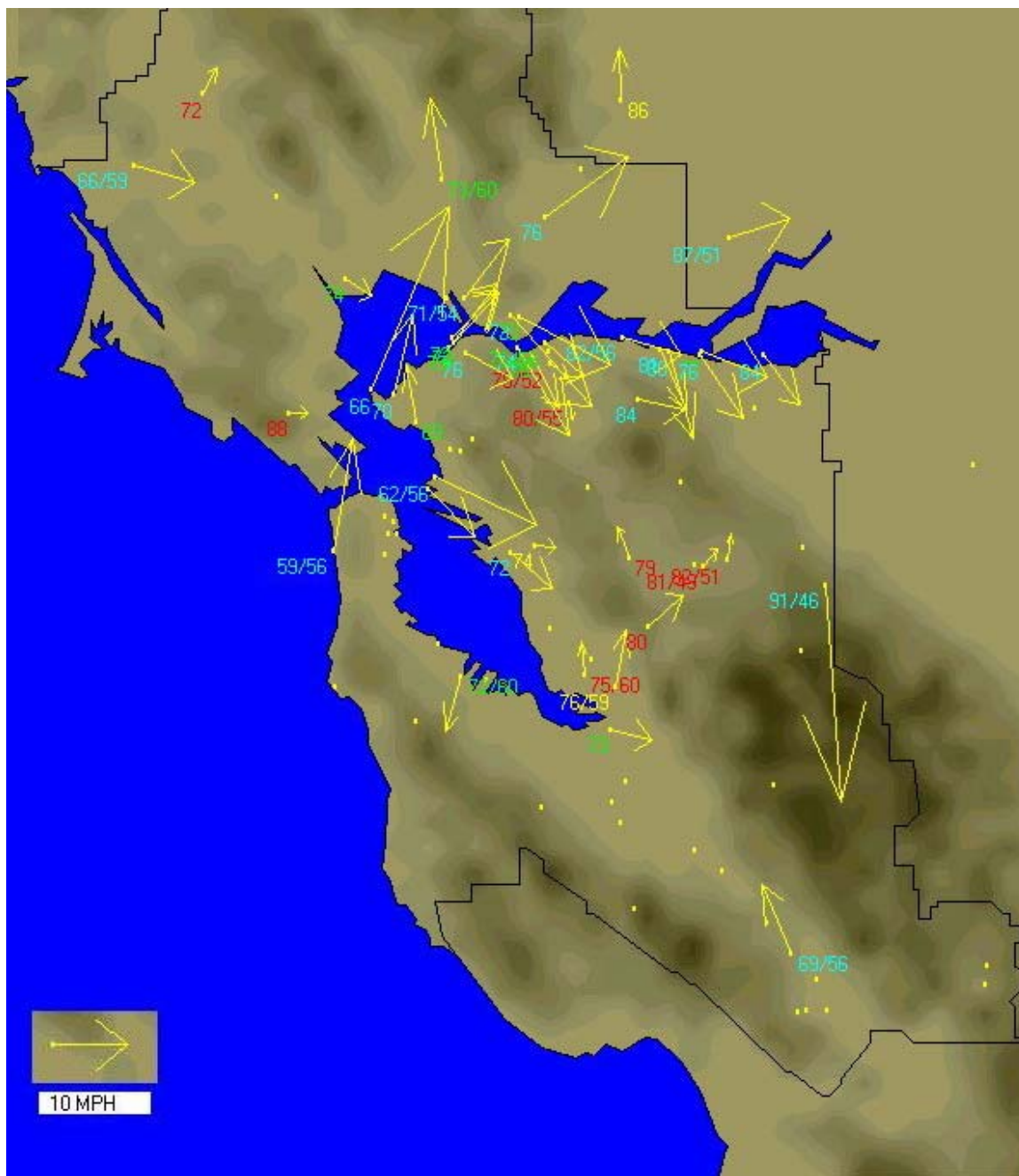
Figure 2-14. Bay area surface-wind observations at 7 AM PST, June 15, 2000. The numbers are temperature and dew point temperature.

Figure 2-15. Bay area surface-wind observations at 2 PM PST, June 15, 2000. The numbers are temperature and dew point temperature.

ENVIRON
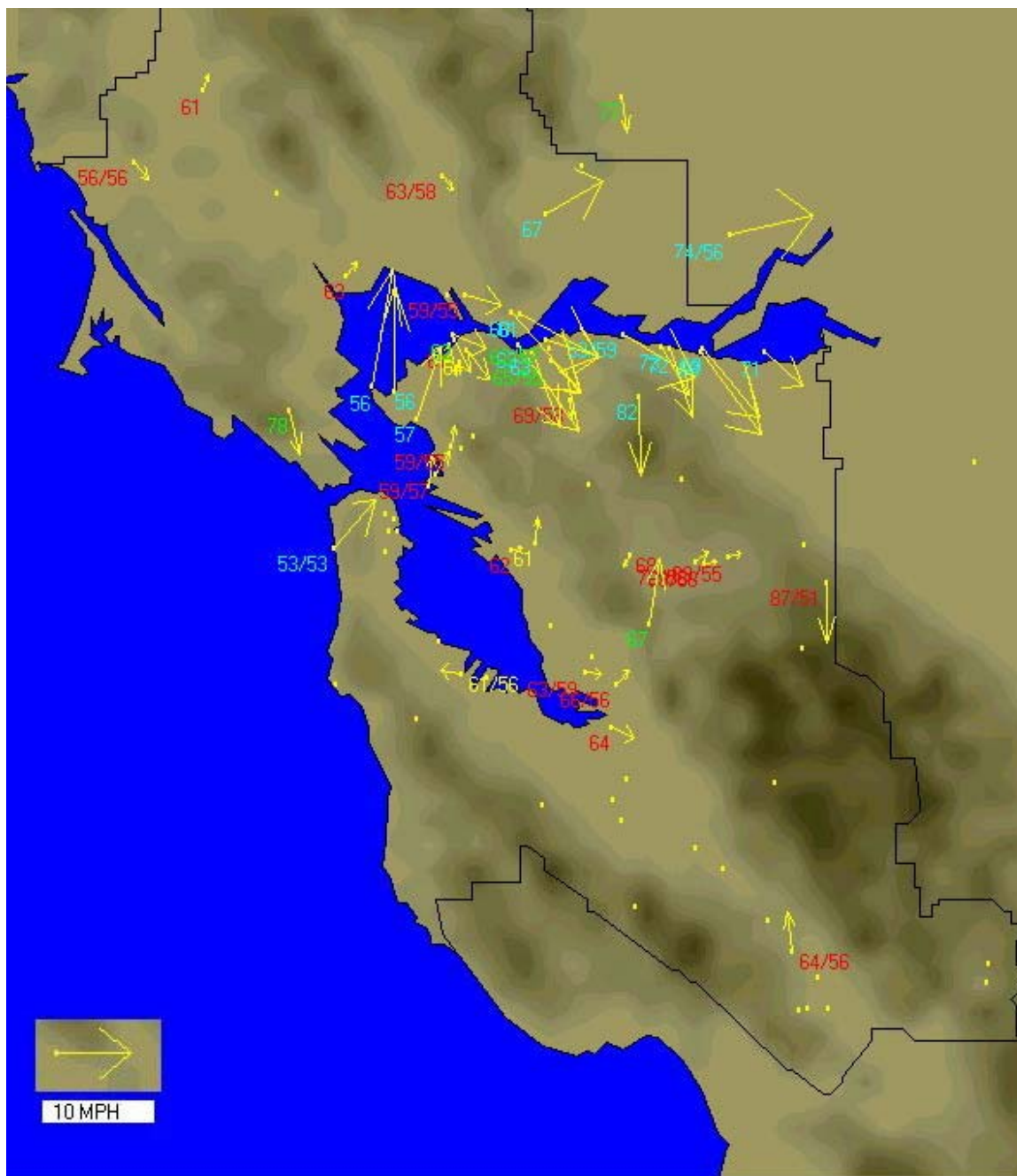
Figure 2-16. Bay area surface-wind observations at 7 AM PST, July 31, 2000. The numbers are temperature and dew point temperature.
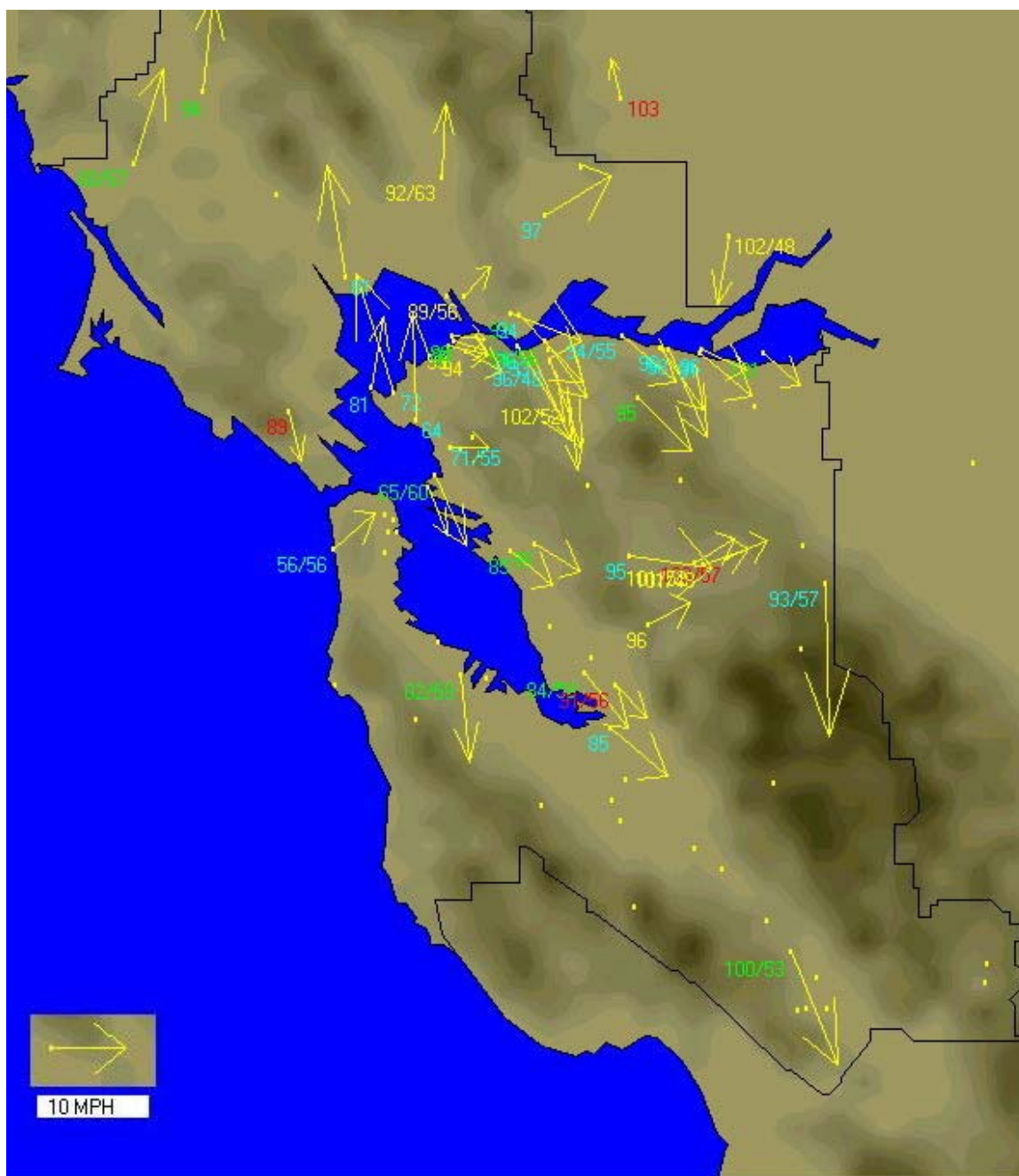
Figure 2-17. Bay area surface-wind observations at 2 PM PST, July 31, 2000. The numbers are temperature and dew point temperature.
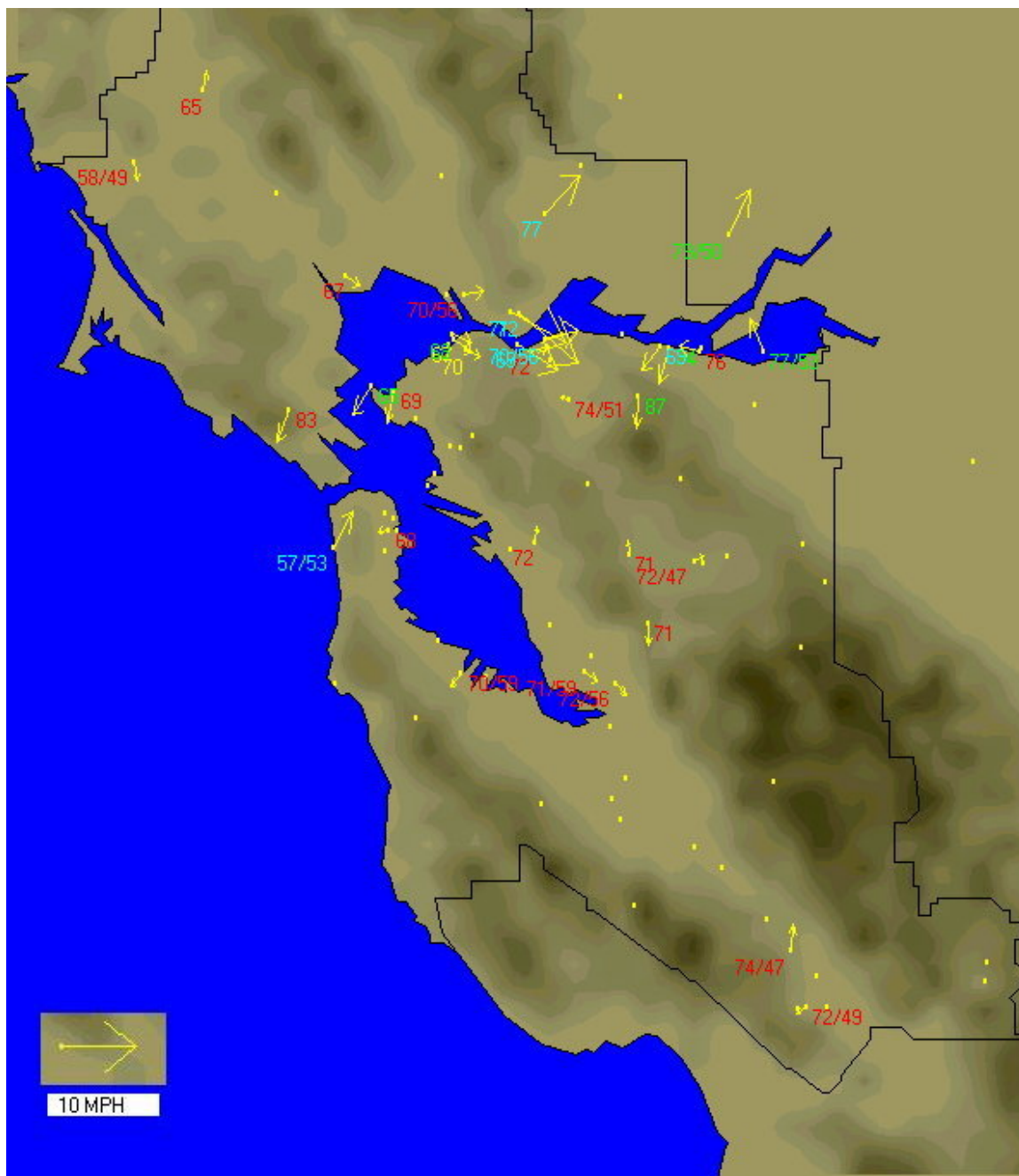
Figure 2-18. Bay area surface-wind observations at 7 AM PST, July 9, 2002. The numbers are temperature and dew point temperature.

Figure 2-19. Bay area surface-wind observations at 2 PM PST, July 9, 2002. The numbers are temperature and dew point temperature.
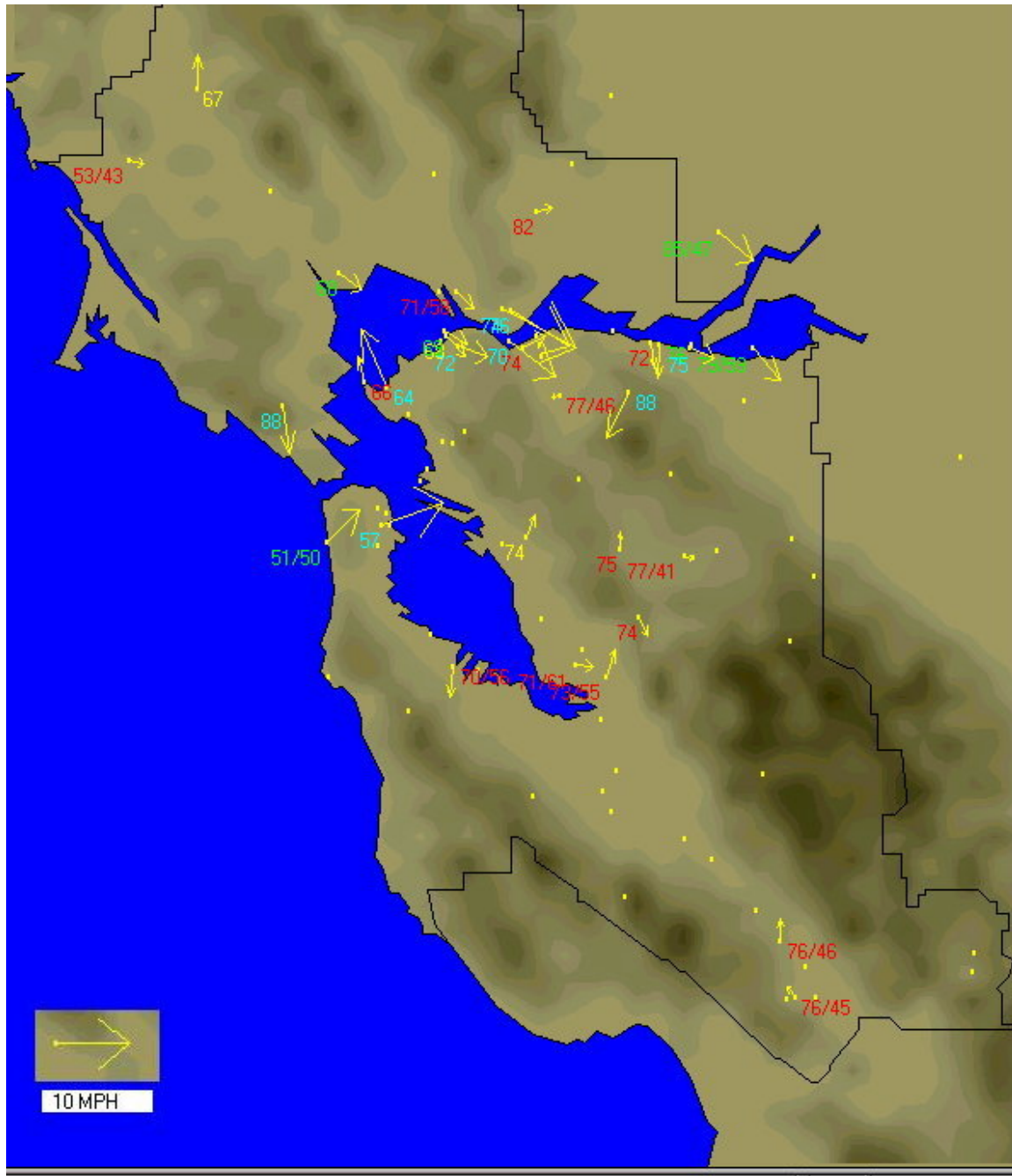
Figure 2-20. Bay area surface-wind observations at 7 AM PST, July 10, 2002. The numbers are temperature and dew point temperature.
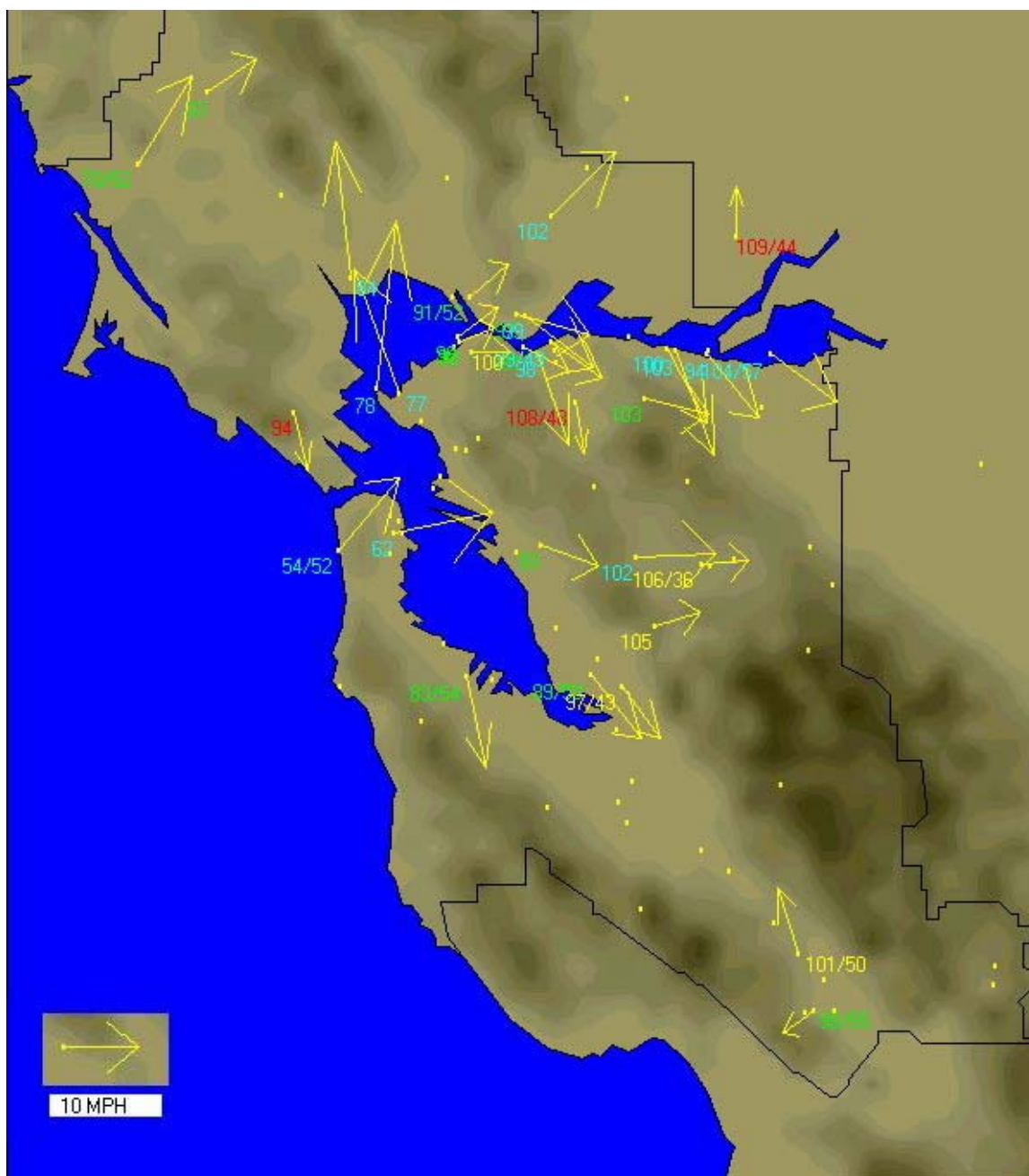
Figure 2-21. Bay area surface-wind observations at 2 PM PST, July 10, 2002. The numbers are temperature and dew point temperature.

other day showed south wind. It poses a challenge to explain the source of ozone precursors at San Martin during high ozone days, especially when the wind is from the south. Perhaps a flow reversal associated with a horizontal shift of the sea breeze front in the southern Santa Clara Valley is a culprit. This front often shifts in response to differing sea/bay breeze strengths between the southern San Francisco Bay and the northern Monterey Bay.

An Evaluation of Transport Potential from the Bay Area

The airborne transport of pollutants is of continuing interest because it affects every region and air basin. This SIP modeling effort is focused on attainment planning for the San Francisco Bay Area, but will provide some information on transport patterns during the selected ozone episodes. This section considers transport of pollutants from the Bay Area to selected sites in the Central Valley and Monterey area for four candidate episode periods (June 11-12, 1999, June 15 and July 31, 2000 and July 9-10, 2002).

Pollutant transport potential between two areas can be assessed by back trajectory analysis, where simulated particles are released at specified times and locations and are transported by winds back in time. The path that the particle takes defines the back trajectory and defines a transport connection between any two points on the back trajectory. For this analysis back trajectories were computed on a NOAA Air Resources Laboratory web site using the HYSPLIT software. However, we must recognize the limitations of this analysis. The back trajectories computed by HYSPLIT used meteorological data with a 40 km grid spacing, which does not resolve wind variations due to narrow mountains, valleys, and gaps in and around the Coastal Ranges. Hence, we have more confidence in the accuracy of these back trajectories in the Central Valley where the terrain is relatively smooth than in the San Francisco Bay and Monterey areas where the terrain is more rugged. The back trajectories in this report were selected as follows:

1. For each of the four candidate episodes, we expanded the date range by 2 days before and 2 days after the episode period. The SIP modeling work is likely to include these additional dates, and thus, we may be able to broaden the analysis of transport.
2. For each expanded episode day, we identified all stations in the Central Valley and Monterey area with ozone exceedances. There were no ozone exceedances in the Monterey area in any of the episode periods. The highest ozone in the Monterey area was 115 ppm observed at Pinnacles at 1700 PST on July 10, 2002. The back trajectory from this Pinnacles observation station is included.
3. For each of these stations, we identified the hour of maximum ozone. These station locations and times defined the initial points for each back trajectory.
4. For all of these initial points, 24- and 48-hour back trajectories were computed.

Figures 2-22 through 2-25 show the composites of all 24-hour back trajectories for each of the four candidate episodes and these are discussed for each episode below.

Transport for July 11-12, 1999

Figure 2-22 shows 17 back trajectories for the 1999 episode. The Sacramento area has definite Bay Area transport connections. Eight of the back trajectories from the Sacramento area passed through the North Bay Area. As noted earlier, the 40 km resolution wind data may not sufficiently resolve the detailed wind variations generated by important topographic features, such as the Carquinez Strait, an ostensible transport corridor. In reality, such features may cause the true back trajectories to be located south of those indicated in Figure 2-22, putting them over the densely populated North Bay Area. Three additional back trajectories ending just north of Sacramento did not reach the Bay Area within 24 hours. Two of these eventually traversed the North Bay Area 24- to 48-hours earlier. The influence of any Bay Area emissions on the Sacramento area will be diluted significantly after traverse times greater than 24 hours.

The six San Joaquin Valley 24-hour back trajectories all had a northwest to southeast orientation. One back trajectory from the Modesto area traversed the Bay Area. The other Modesto area back trajectory and the two Fresno back trajectories traversed the Stockton area 24-hours earlier and traversed the Bay Area 24- to 48-hours earlier. The back trajectory from Bakersfield and one from near the Sequoia National Park did not show any Bay Area connections.

Transport for June 15, 2000

Figure 2-23 shows six back trajectories for the June 15, 2000 episode. All back trajectories ended in the Sierra. There were no Bay Area connections even up to 48 hours.

Transport for July 31, 2000

Figure 2-24 shows five back trajectories for the July 31, 2000 episode. The Fresno back trajectory traversed the Bay Area. The two 24-hour Bakersfield back trajectories did not traverse the Bay Area but the 48-hour back trajectories did. The two back trajectories from the Sacramento area traversed the North Bay Area.

Transport for July 9-10, 2002

Figure 2-25 shows 19 back trajectories for the July 9-10, 2002 episode. Three of the four Sacramento area back trajectories traversed the North Bay Area. The fourth one had no Bay Area connection. The three back trajectories from Merced County had clear Bay Area connections, with two passing through the central Bay Area and the other passing through the North Bay Area. Of the eight back trajectories from Fresno, only one traversed the Bay Area within 24 hours, and two others traversed the North Bay Area within 48-hours. The three back trajectories from Bakersfield passed over Fresno within 24 hours and two of these reached the central Bay Area within 48-hours. The back trajectory from the Pinnacles meandered through the Santa Clara Valley and East Bay Area before reaching the North Bay Area 24 hours later.
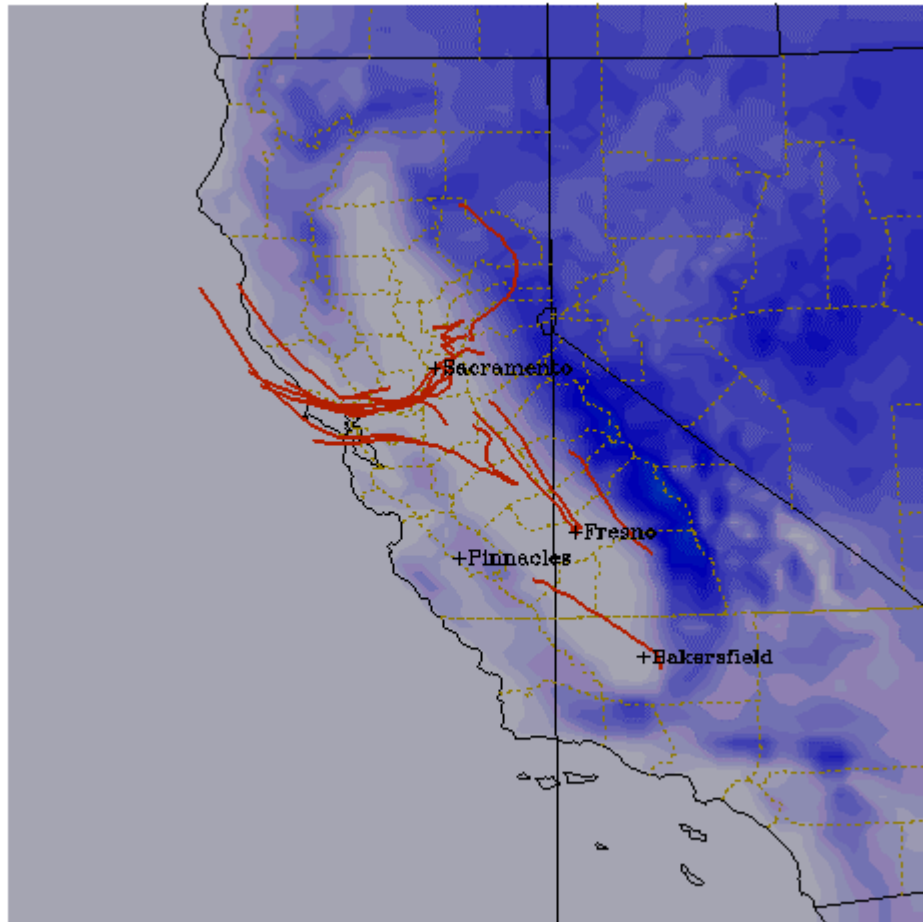
Figure 2-22. The 24-hour back trajectories for the July 11-12, 1999 episode.
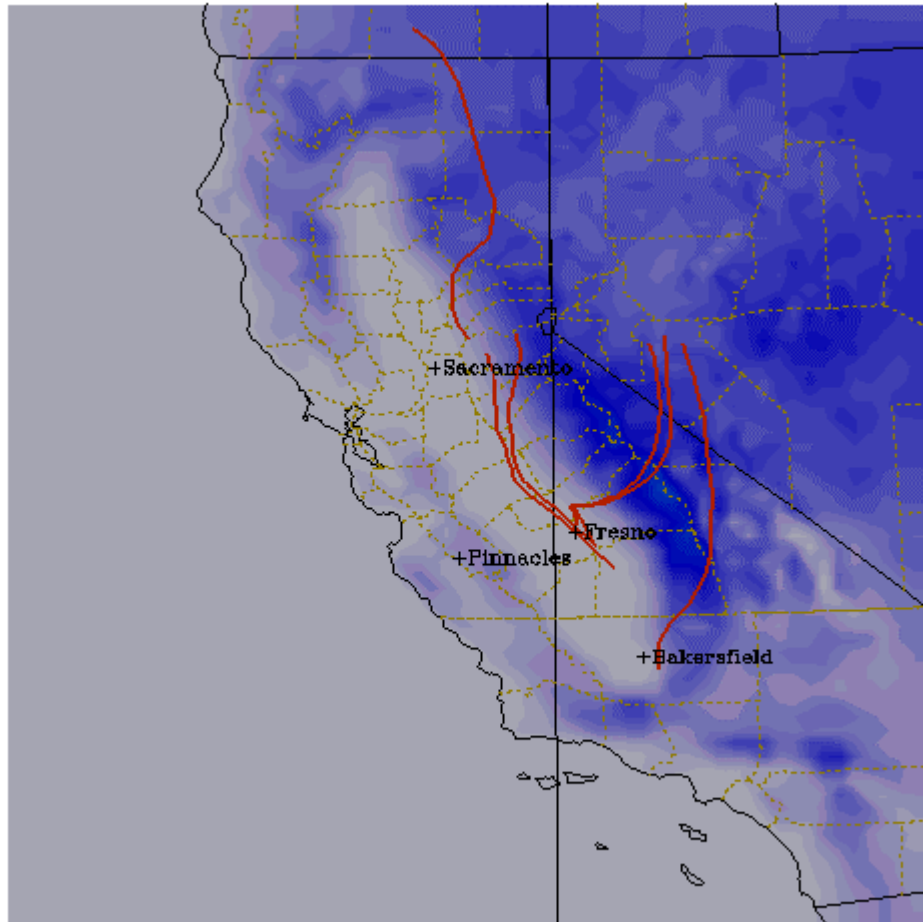
Figure 2-23. The 24-hour back trajectories for the June 16, 2000 episode.
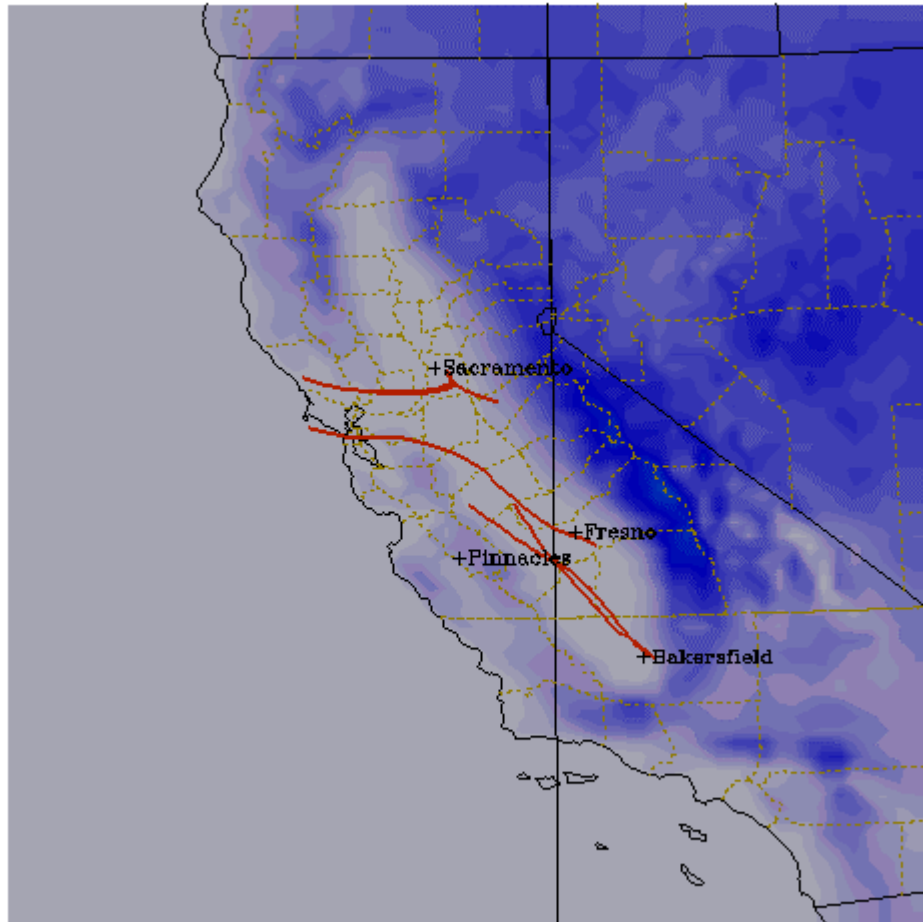
Figure 2-24. The 24-hour back trajectories for the July 31, 2000 episode.
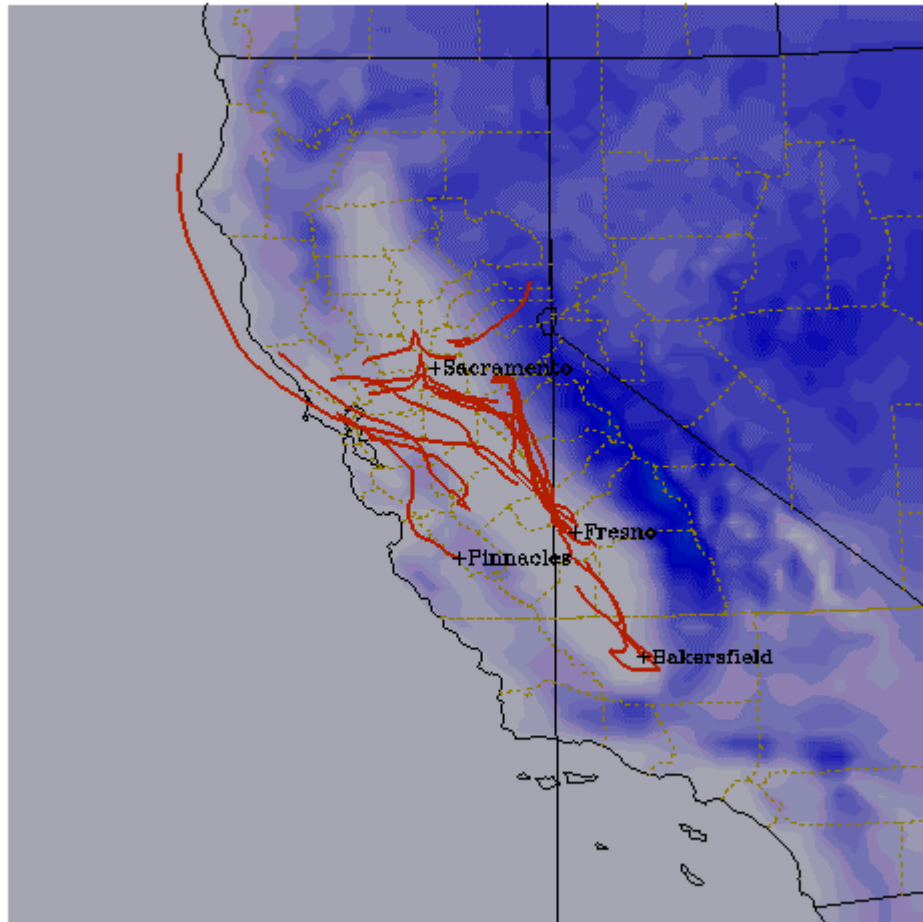
Figure 2-25. The 24-hour back trajectories for the July 9-10, 2002 episode.

Summary

The potential for transport of pollutants from the Bay Area to the Central Valley and Monterey areas for four candidate Bay Area SIP episodes were assessed using the HYSPLIT back trajectory analysis. The results are summarized in Table 2-5. More than 50% of the Sacramento Valley 24-hour back trajectories pass through the Bay Area. For San Joaquin Valley, this ratio is less than 25%. This is understandable because most high ozone days occur during stagnant or weak wind conditions. Hence, pollutants from the Bay Area will not be able to reach Fresno or Bakersfield within 24 hours. If the back trajectory computation is extended to 48 hours, this ratio increases to 65% in the Sacramento Valley and to 55% in the San Joaquin Valley. The one back trajectory computed for the Monterey area does show a clear Bay Area transport connection.

Table 2-5. The ratios of the number of back trajectories passing through the Bay Area to the total number computed. The 24- and 48-hour columns indicate the ratios for the 24- and 48-hour back trajectories, respectively. Sacramento Valley includes Stockton area. San Joaquin Valley includes Merced County area.

Episode	Sacramento Valley		San Joaquin Valley		Monterey Area	
	24-hour	48-hour	24-hour	48-hour	24-hour	48-hour
7/11-12/99	7/10	9/10	2/7	5/7	0/0	0/0
6/15/00	0/1	0/1	0/5	0/5	0/0	0/0
7/31/00	2/2	2/2	1/3	3/3	0/0	0/0
7/9-10/02	3/4	3/4	4/14	8/14	1/1	1/1
Total	9/17	11/17	7/29	16/29	1/1	1/1

We rank these episodes for suitability of use in transport analyses as follows:

1. The July 11-12, 1999 episode.
2. The July 31, 2000 or the July 9-10, 2002 episodes. The July 9-10, 2002 episode includes the only Pinnacles high ozone case.

No back trajectories during the June 15, 2000 episode showed any transport connection from the Bay Area to the Central Valley or the Monterey area.

Discussion and Conclusions

The statistical analysis has looked at representativeness from several points of view and suggests the following points:

1. There has been substantial progress in reducing ozone in the south bay region. In contrast, there has been little progress for eastern sites. Almost all exceedance days include high ozone at eastern sites. Thus, it seems reasonable that any modeled day should include high ozone from eastern sites.

2. The cluster analysis suggests that the days with high ozone at Fremont (6/24/95 and 7/15/95) are substantially different from other exceedance days. This, coupled with the fact that the emissions leading to high ozone at Fremont may well be different from those at eastern sites, and the discussion in (1) suggests that these days be excluded from modeling and be considered a category that is not represented in the modeling analysis.

3. High ozone occurs at eastern sites on both weekdays and weekends. The emissions patterns will clearly differ for these two periods. Thus, ideally, modeling should include both weekend and weekday exceedances.

4. The cluster analysis found two main categories of clusters: days with exceedances at isolated individual sites where the rest of the District was relatively clean; and days with exceedances at several sites and regions. Cluster 2 days have much lower mean ozone than Cluster 1 and differ with respect to several meteorological variables. In particular, Cluster 2 days are cooler in the south bay and have somewhat stronger winds through the Carquinez Strait and Santa Clara Valley. In addition, Cluster 2 afternoon 850 mb winds almost all contained a westerly component, whereas over half the Cluster 1 winds contain an easterly component. It is not unreasonable to assume the dynamics of ozone formation on these two types of days could be different. Thus, it seems reasonable to plan to model days from within both of these categories.

5. Both potential CCOS modeling days fell into Cluster 2, the cluster representing an isolated high ozone event. Thus, these days may be acceptable to model Cluster 2 days, but another episode is necessary to cover Cluster 1 days. Also, since both CCOS days fell on weekdays, it would be valuable if the additional episode covered at least one weekend day, and the CCOS days represented single day episodes, so it would be valuable to include a multi-day episode.

The meteorological analysis of the 6 candidate ozone episode days identified from the statistical representativeness evaluation leads to the following conclusions:

1. The weather patterns of the 6 ozone episode days are similar, with high 500 mb contour heights, high 850 mb temperatures and low inversion layers.

2. The 2 PM PST surface wind patterns are similar on all 6 days in all regions except in the Santa Clara Valley. One source area of ozone precursors at Livermore is likely the Martinez area, where northwest winds prevail. Another source area of ozone precursors at Livermore may be the area surrounding the San Francisco Bay. Precursors from this area arrive via the sea breeze through the Castro Valley-Dublin gap, which starts after 10 AM.

3. The afternoon wind patterns at Livermore are similar for all 6 ozone episode days. The mechanism for ozone production at Livermore in Clusters 1 and 2 may also be similar. Based on this analysis, it would appear that any of these 6 days are as good as another for modeling ozone at Livermore.

4. There is no clear relation between the wind direction and the observed ozone at San Martin. Therefore, it poses a challenge to determine the source of ozone precursor at San Martin during high ozone days, especially during south wind days.

Selected Episodes for Modeling

The table below summarizes the six episode days described above.

Episode Dates	Days of Week	Peak ozone (ppb)	# Exceedances	Cluster Category	PM Wind Pattern
July 11-12, 1999	Sun, Mon	156 Concord	6	1	Similar
June 15, 2000	Thurs	152 Livermore	1	2	Similar
July 31, 2000	Mon	126 Livermore	1	2	Similar
July 9-10, 2002	Tues, Wed	160 Livermore	2	1	Similar

Based upon the review above, and the criteria for data availability, we propose to model four exceedance days for the 2004 SIP sumbittal, in the following order:

1) July 31, 2000
2) June 15, 2000
3) July 11 and 12, 1999.

The June and July 2000 days occurred during the Central California Ozone Study (CCOS), a field monitoring program that collected extensive meteorological and aerometric measurements for use in the analysis and the modeling of ozone throughout central California. Both of the 2000 days fell into one of the two main episode categories described above.

The 1999 days represent the other frequently occurring ozone pattern category. July 11 was a Sunday and July 12 was a Monday, which should satisfy the need to evaluate weekend-weekday issues. Data for this period is quality assured and archived (data for July 9-10, 2002 is not readily available and this unavailability may cause delays in our SIP schedule). Also, this episode experienced more wide-spread exceedances than other periods (3 per day).

The CARB will provide emission estimates for the year 2000 (as described later in Section 5); hence, the adjustment for estimating 1999 emissions will not be as large as required for earlier years.

DESCRIPTION OF EPISODES

Histories of daily maximum observed ozone for inside and outside the SFBA for the 1999 and 2000 seasons are presented in Figures 2-26 and 2-27, respectively. Figure 2-26 shows that the largest SFBA exceedance occurred on 7/12/1999. Exceedances over this season were infrequent and peak values varied significantly from the generally clean levels around 60 ppb. The ozone observations outside the SFBA were consistently around the 120 ppb level.

Figure 2-27 shows that the SFBA exceedances (6/15/2000 and 7/31/2000) also occurred infrequently and that the exceedance values varied significantly from the generally clean (60 ppb) 2000 season. During this season the daily maxima inside and outside the SFBA were better correlated than they were during the 1999 season.
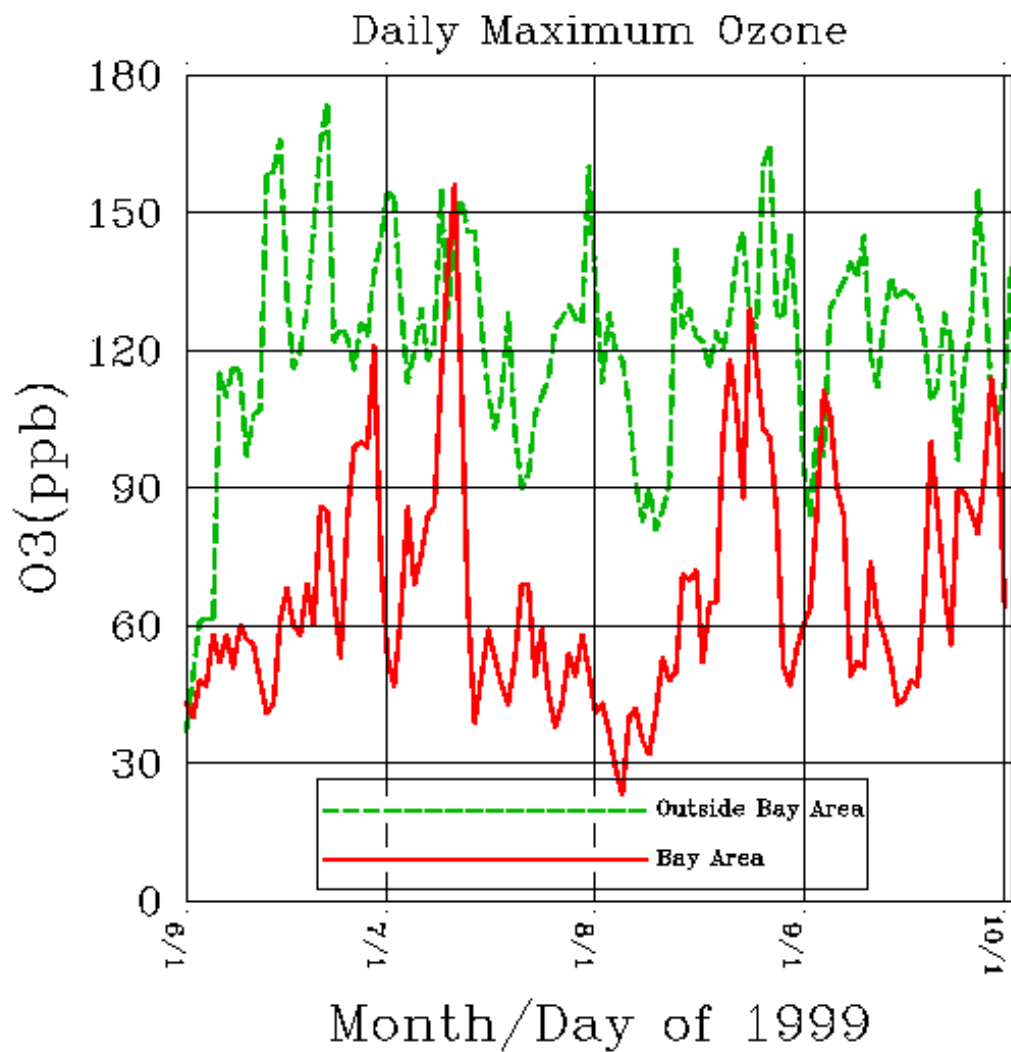
Figure 2-26. Daily maximum observed ozone from 6/1/1999 to 10/1/1999.
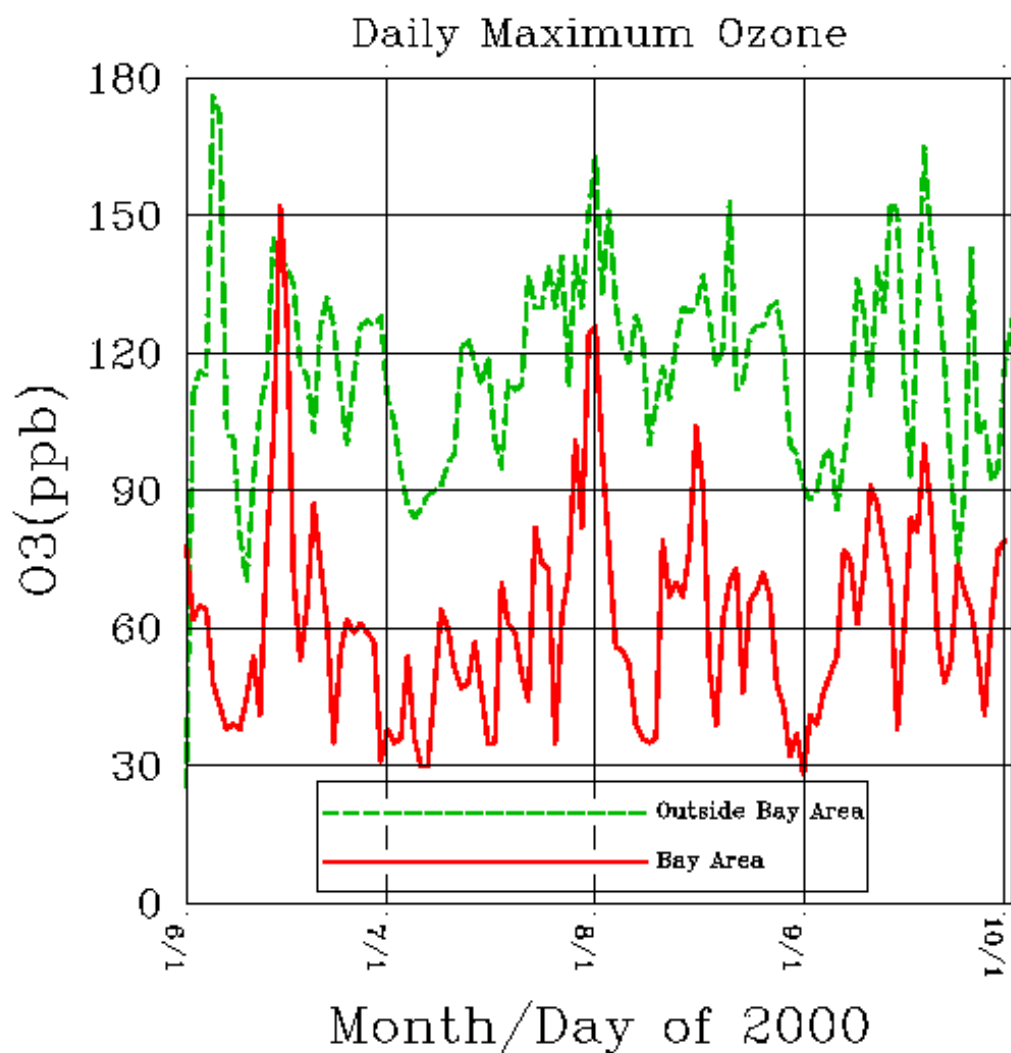
Figure 2-27. Daily maximum observed ozone from 6/1/2000 to 10/1/2000.

Figures 2-28 through 2-31 show the spatial distribution of daily maximum ozone for each of the four selected exceedance modeling days. On 7/11/1999 the exceedances were wide-spread and include Sacramento, Fresno, and Los Angeles areas. On 7/12/1999 the high ozone values outside the SFBA diminished from the previous day, leaving the highest ozone observations in the SFBA. On 6/15/2000 isolated high ozone values were quite localized at Livermore, while clean conditions existed around Sacramento, and moderate values existed around Fresno. On 7/31/2000 ozone values just above the clean air standard were observed in Livermore and Fresno, while moderate values were observed around Sacramento.

Figures 2-32 through 2-34 present hourly ozone time series ("histories") for the SFBA sites that measured exceedances during these four episode days. Figure 2-32 presents ozone time series for 7/11-12/1999 at the Livermore (Old First Street) and Concord sites. The ozone observations were similar for these two locations indicating that high levels of ozone were widespread over the East Bay. Figure 2-33 presents the time series at the two Livermore sites for 6/15/2000. These sites were approximately one mile from each other, and so their time series were quite similar; no other sites measured ozone exceedances. Figure 2-34 presents the ozone time series at the two Livermore sites for 7/31/2000l; these were similar (but lower) to 6/15/2000.
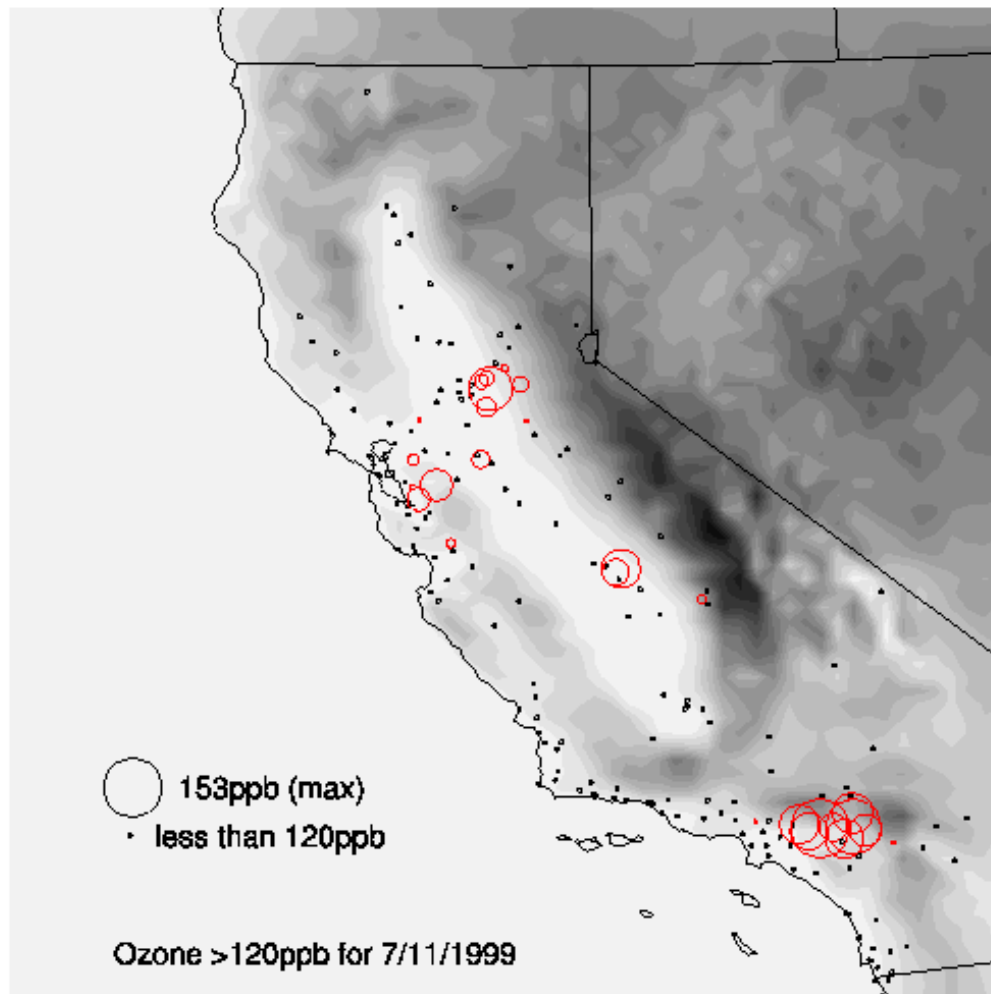
Figure 2-28. Spatial distribution of daily maximum ozone observations greater than 120 ppb (red) for 7/11/1999.
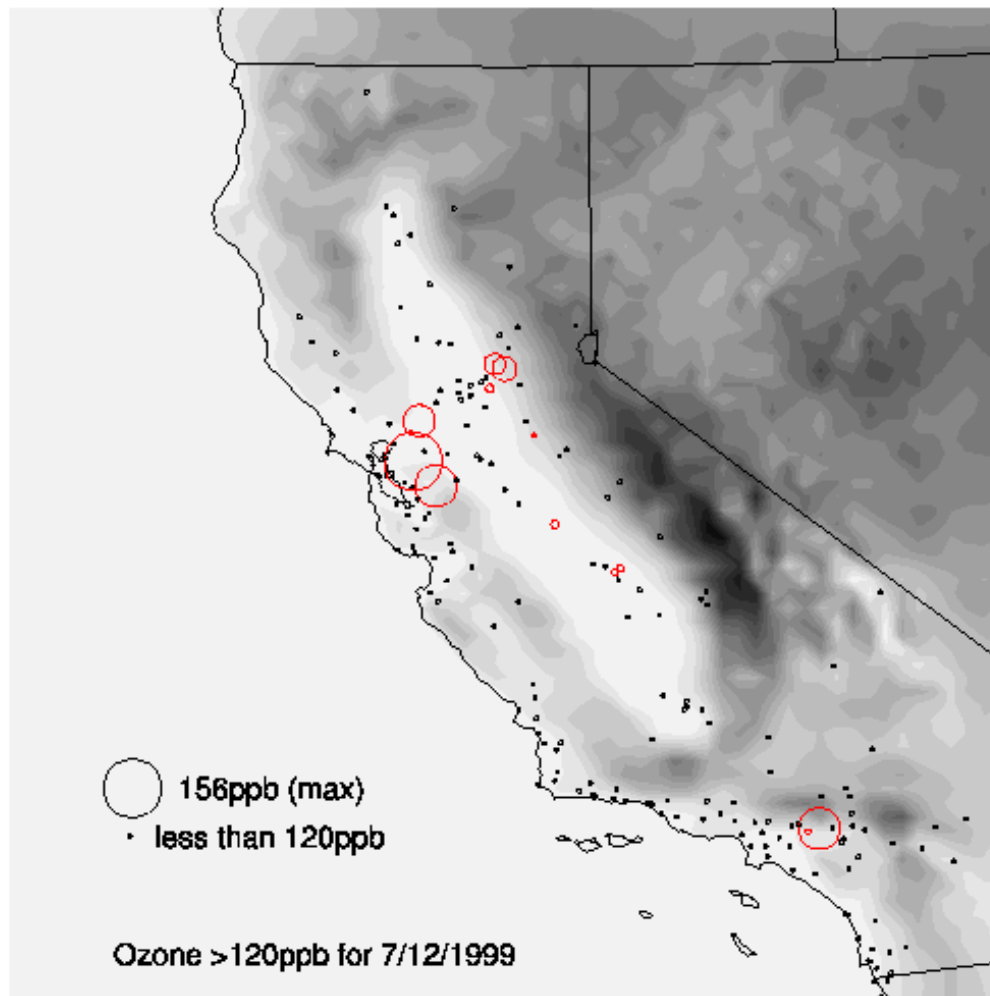
Figure 2-29. Spatial distribution of daily maximum ozone observations greater than 120 ppb (red) for 7/12/1999.

December 2002

ENVIRON
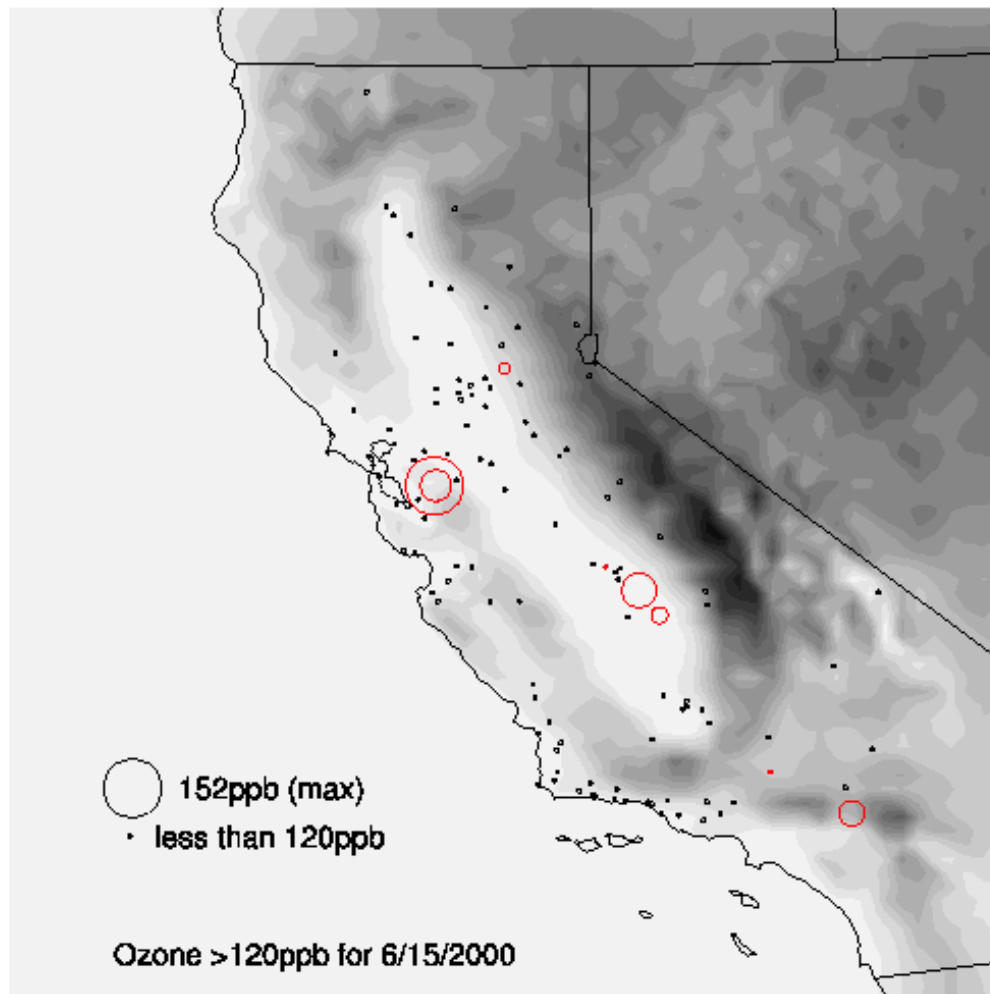
Figure 2-30. Spatial distribution of daily maximum ozone observations greater than 120 ppb (red) for 6/15/2000.

H:\BAAQMD\Modeling Protocol\Draft 4\Section_2.doc

2-43
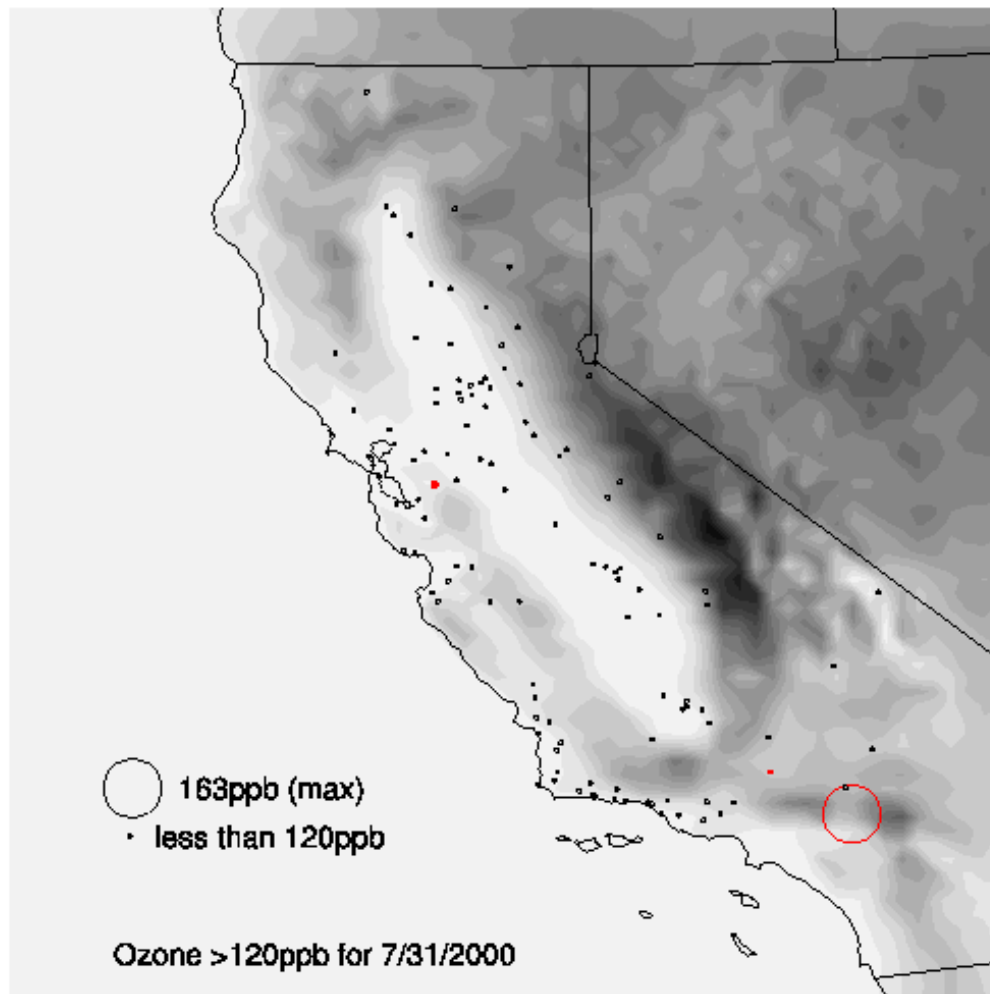
Figure 2-31. Spatial distribution of daily maximum ozone observations greater than 120 ppb (red) for 7/31/2000.

Ozone History for 7/11—12/1999
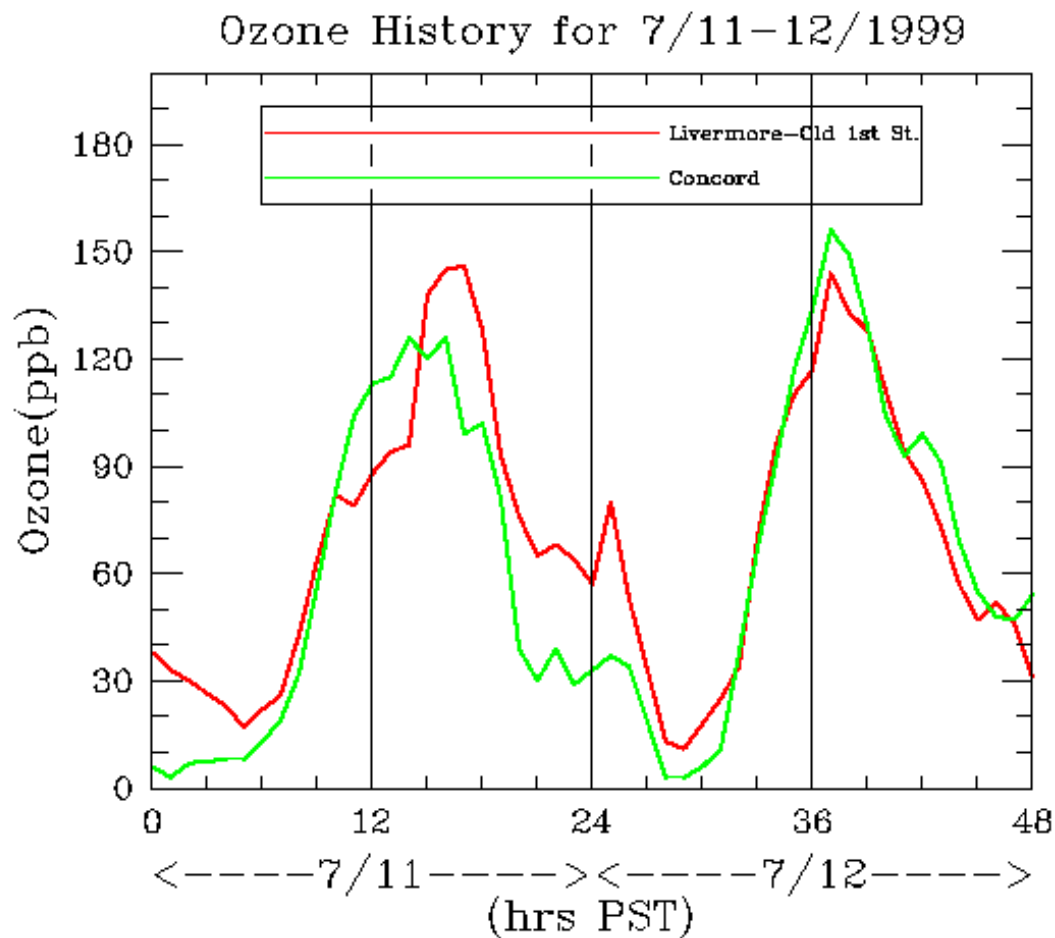
Figure 2-32. Ozone time series at the two SFBA stations with the highest ozone observations during 7/11-12/1999.
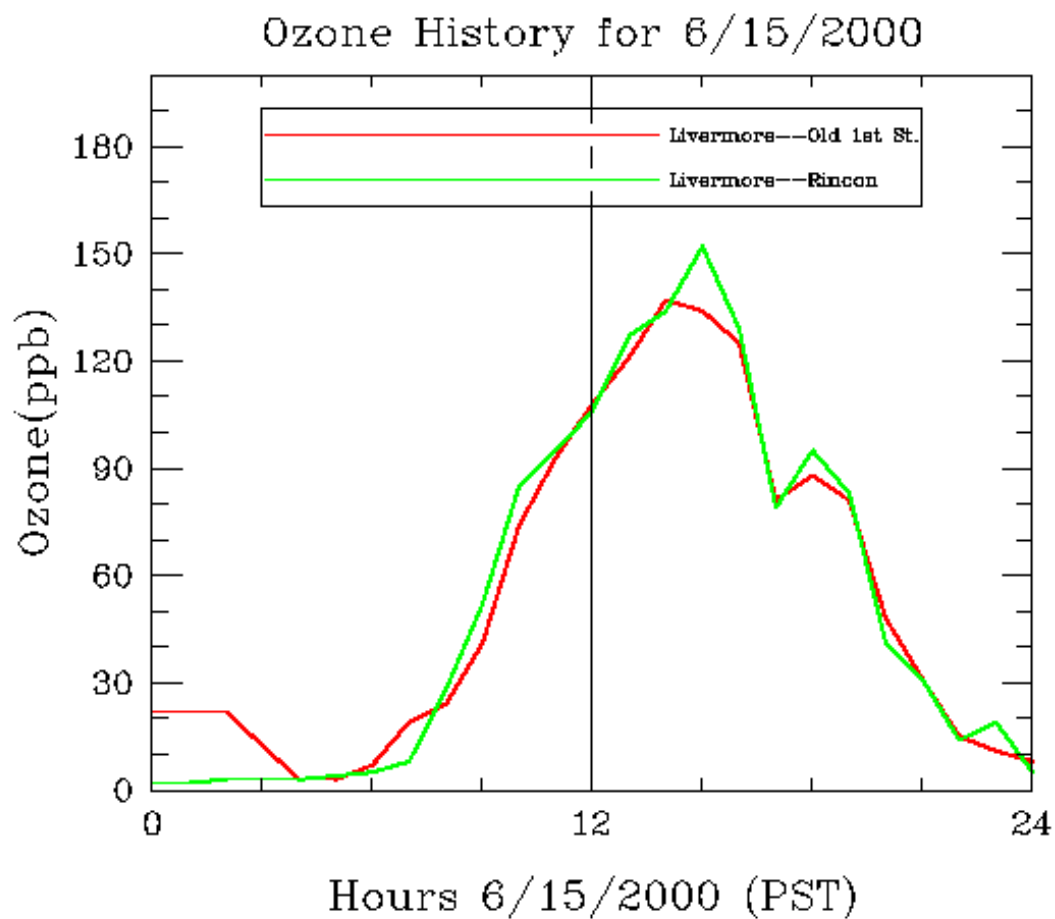
Figure 2-33. Ozone time series at the two SFB stations with the highest ozone observations during 6/15/2000.
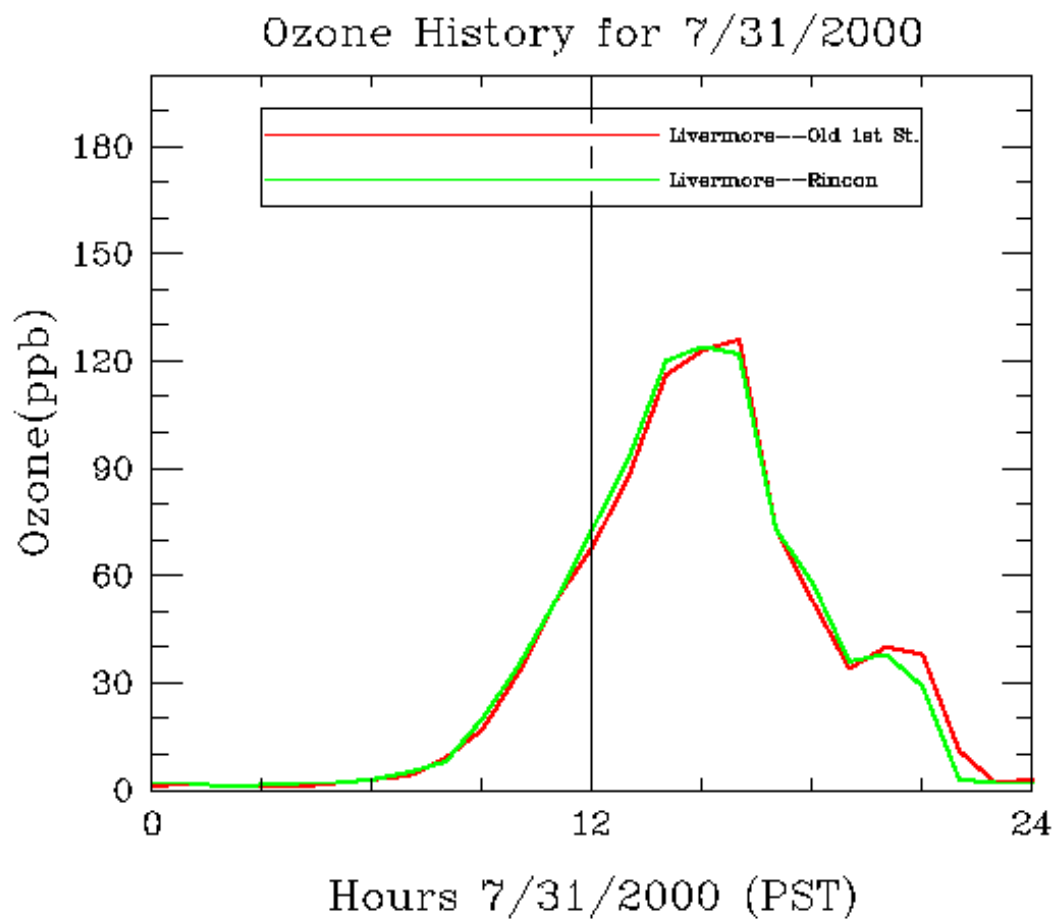
Figure 2-34. Ozone time series at the two SFB stations with the highest ozone observations during 7/31/2000.

Summary of Meteorology

Lehrman et al. (2001) describe the CCOS meteorological conditions and their relationship to ozone values:

"The relationship between the dispersion of ozone and ozone precursors in California and large-scale synoptic weather patterns is well known. During the summer ozone season, the extension of the eastern Pacific high over the western US effectively blocks the influx of cyclonic weather systems into California from the Gulf of Alaska, and allows the entrenchment of large static air masses which are typically warm, stable, and poorly mixed. The strength and persistence of the resultant boundary layer mixing and transport patterns affects the magnitude and duration of ozone events in Central California. High-pressure ridges and low-pressure troughs in the mid to upper atmosphere are particularly efficient indicators of ozone formation conditions. … Two synoptic scale meteorological parameters, which historically have correlated well with ozone formation and fate in California, are the height of the 500 mb surface and the temperature at the 850 mb level. The time history of 500 mb heights at a fixed location is a general indicator of the behavior of the 500 mb surface indicating pressure ridges and troughs. The 850 mb temperature is a measure of large-scale subsidence, which produces stable layers in the atmosphere and limits vertical dispersion of ozone and precursors."

Figures 2-35 and 2-36 show the variation of the 850 mb temperature and 500 mb heights at Oakland for the 1999 ozone season from 6/1/1999 to 9/30/1999. The 500 mb height at the beginning of the season dipped to as low as 5470 m on 6/3/1999. It increased rapidly and varied over a much narrower range over the entire ozone season afterwards. The few days leading to 7/11/1999 were characterized by the gradual building of the 500 mb heights from 5760 m on 7/4/1999 to a high of 5950 m on 7/10/1999. The 500 mb heights stayed at 5940 for 7/11-12/1999 and there were closed height contours over Northern California over this two-day period. The building of the 500 mb height can be easily seen on the weather maps (not shown), where the 5880 contour line moved from Central California to Washington State during this period of time. The 850 mb temperature increased from 9.6 $^{\circ}$C on 7/4/1999 to a high of 27.2 $^{\circ}$C on 7/12/1999 during the building of the 500 mb height. The peak 850 mb temperature was reached two days after the peak 500 mb height and this can be easily explained by the continued warming from the downward motion in a high pressure area. The highest surface temperature reached 113°F on 7/11/1999 and 115°F on 7/12/1999, both at Redding. The highest temperature in the San Joaquin Valley was 106°F at Fresno for both days.

Figures 2-37 and 2-38 show the variation of the 850 mb temperature and 500 mb heights at Oakland for the 2000 ozone season from 6/1/2000 to 9/30/2000. The peak 500 mb height and the peak 850 mb temperature correlated well with the high SFBA ozone observations during this season. Lehrman et al., 2001, describe the synoptic conditions leading up to 6/15/2000:
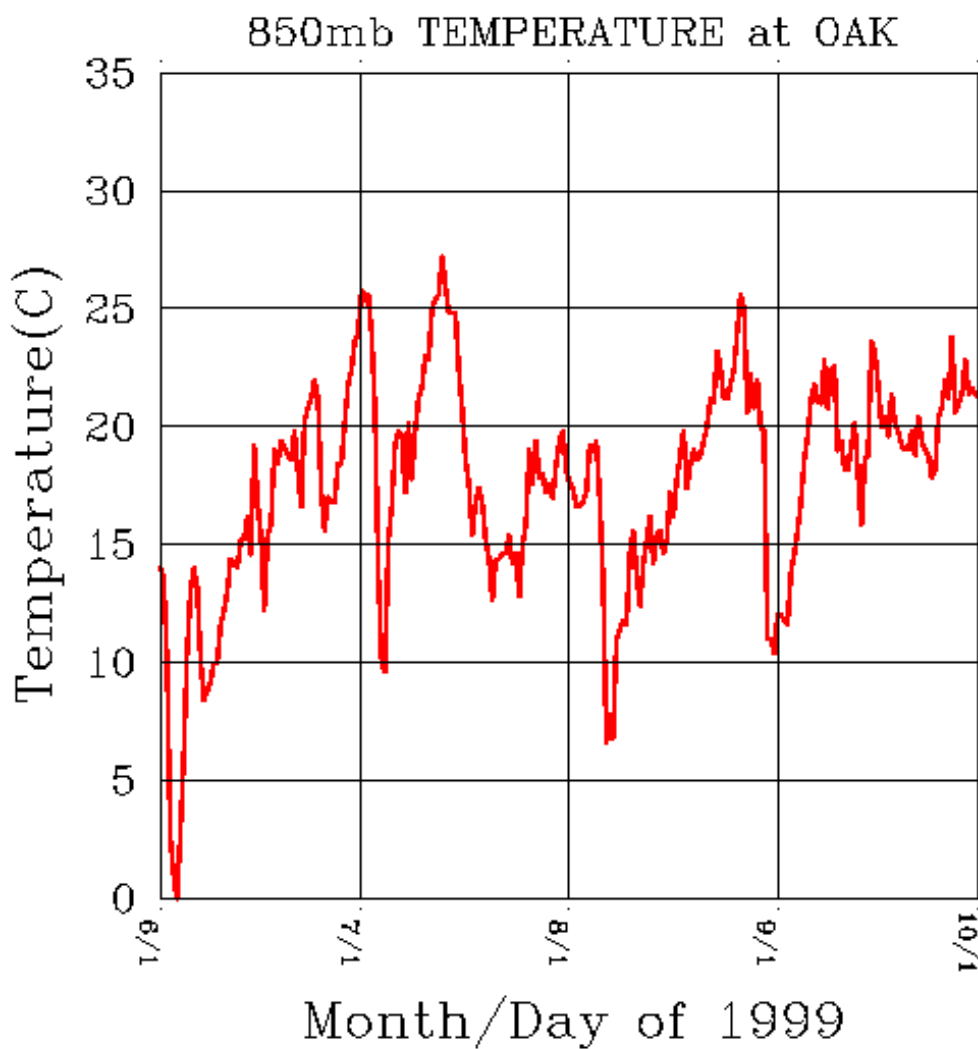
Figure 2-35. 850 mb temperatures at Oakland from 6/1/1999 to 10/1/1999.
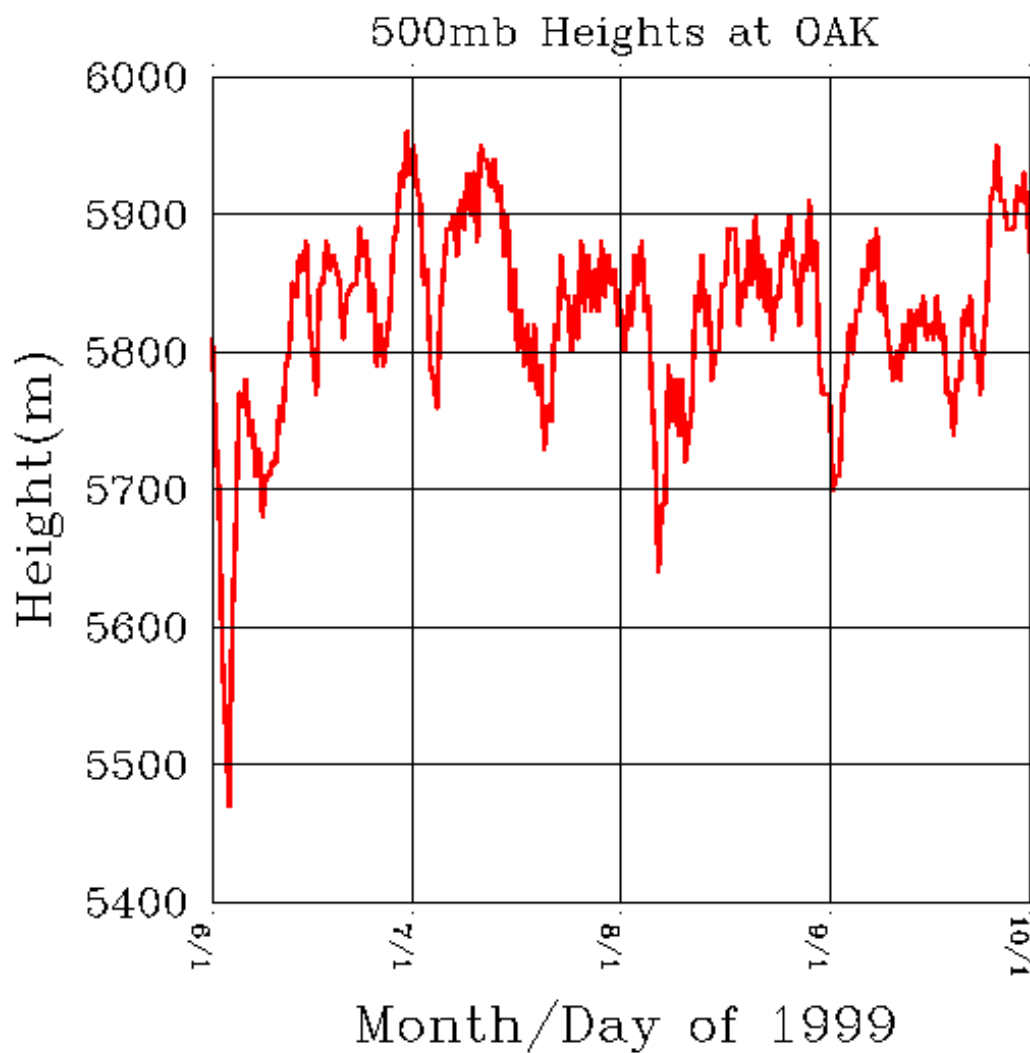
Figure 2-36. 500 mb heights at Oakland from 6/1/1999 to 10/1/1999.
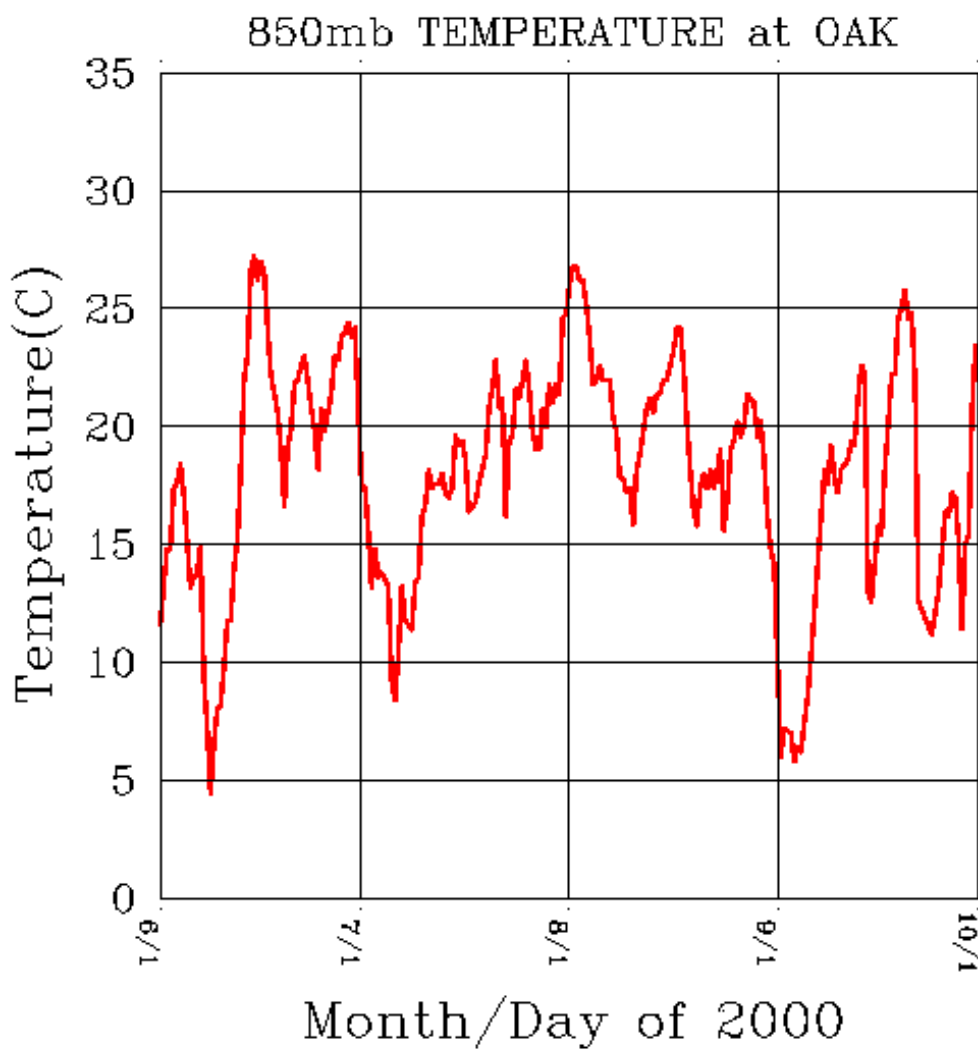
Figure 2-37. 850mb temperatures at Oakland from 6/1/2000 to 10/1/2000.
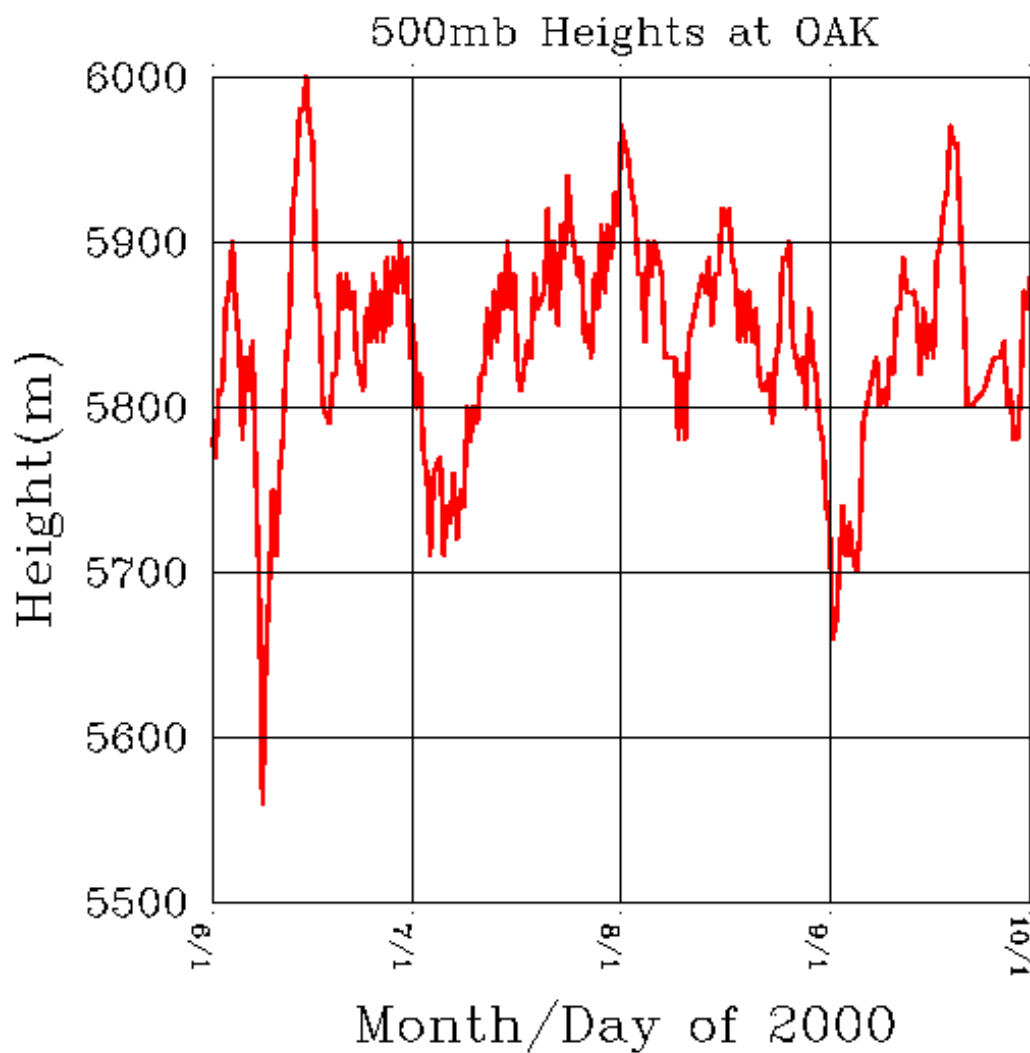
Figure 2-38. 500 mb heights at Oakland from 6/1/2000 to 10/1/2000.

"The OAK (Oakland) 500 mb height had increased from a low of 5,650 m on June 8 to a maximum of 6,000 m on June 14. During that same period, the OAK 850 mb temperature increased from 7°C on June 8 to a high of 27°C on June 14. As the ridge progressed towards the east-southeast, flow aloft remained from the north throughout the period. This slowly encouraged the onset of offshore flow across the project area during that time. Ozone concentrations increased steadily as the ridge approached with peak ozone values in excess of the Federal and State Standards…"

They also describe conditions for the 7/31/2000 episode:

"By July 25, the ridge had weakened slightly and dropped southeastward into eastern New Mexico and a trough developed along the West Coast from Point Conception to British Columbia. This resulted in the lowering of 500 mb heights and 850 mb temperatures somewhat during July 25 and 26. However, on the 27th, the high-pressure ridge once again regressed towards the west and strengthened somewhat to become centered once again in the Four Corners area. With this regression of the ridge, the 850 mb temperature and 500 mb heights at OAK once again rose during that period and continued to rise through July 30 … During the IOP of July 30 through August 2, the ridge remained strong and continued to slowly regress towards the west until it was centered near Reno, Nevada by July 31. The OAK 850 mb temperature during the IOP reached as high as 27°C and the 500 mb height topped at 5,970 m … Elevated ozone concentrations persisted in the project area for several days after the IOP, which ended on August 2."

DATABASES FOR THE EPISODES

Data for the 1999 Episode

Data to support the modeling and analysis of the 1999 episode will be from routine sources. During summer 1999, the CARB and the AQMDs operated over 150 surface-based air quality monitoring stations throughout northern and central California. Many of these sites routinely measured O$_3$, NO$_x$, CO and hydrocarbons. Existing PM10 measurements acquired filter samples every sixth day. A few of the PM10 sites had continuous monitors that measured hourly PM10 everyday. A few routine PM2.5 measurements sites were also in operation. Districts in the Sacramento and San Joaquin Valleys are required to routinely operate photochemical assessment monitoring stations (PAMS) as part of their State Implementation Plans. Each PAMS station measures speciated hydrocarbons and carbonyl compounds, O$_3$, NO$_x$, and surface meteorological data. Additionally, each area must monitor upper-air meteorology at one representative site.

An extensive but uncoordinated network of surface meteorological monitoring sites is routinely operated by the CARB, BAAQMD, SJVAPCD, SMAQMD, the National Oceanic and Atmospheric Administration (NOAA), the California Irrigation Management Information Service (CIMIS), Interagency Monitoring of PROtected Visual Environments (IMPROVE), the National Weather Service (NWS), Pacific Gas and Electric Company (PG&E), the U.S. Coast Guard, Remote Automated Weather Stations (RAWS), and a few additional agencies. Wind speed and direction, temperature, and relative humidity are the most common measurements. Surface pressure and solar radiation measurements are also common. A few

sites measured ultraviolet radiation in the Sacramento and San Joaquin Valleys, and in Santa Barbara County.

The CARB operated two profilers (with RASS) in the San Joaquin Valley, and the San Joaquin Unified APCD and Sacramento Metropolitan AQMD operate one profiler/RASS each as part of their PAMS monitoring program. The SJVAPCD also operated a profiler at Tracy during the 2000 CCOS. Military facilities with operational profilers include Travis AFB, Vandenberg AFB, and the Naval Post Graduate School in Monterey. Radiosonde measurements of winds, temperatures, and humidity aloft are routinely made twice per day at Oakland and, according to military base requirements, at Vandenberg, Edwards, and Pt. Mugu.

Routine measurements of pollutant emissions from stacks are required for large industrial sources, such as utility boilers. Traffic count data were also routinely collected at many freeway locations in central California.

Finally, polar-orbiting and geostationary satellites collected an enormous amount of radiometric data that yield useful products, including the total ozone column, cloud cover, sea surface temperature, vegetative cover, and surface albedo throughout California.

Data for the 2000 Episodes

Two of the 2000 episodes identified as candidates for this study occurred during the Central California Ozone Study (CCOS). Most of the data that will be used for SIP modeling and analysis -- for generating model inputs, for model evaluation, and for corroborative studies -- will therefore be derived from the CCOS database. During CCOS, when high ozone episodes were forecast, an intensive operation period (IOP) was launched and additional special field study data were collected. The 7/31/2000 episode occurred within an IOP and it will therefore benefit from many special field-study observations. However, the 6/15/2000 episode occurred before most of the special study data from CCOS were being gathered, so this episode will rely primarily on routine data within the CCOS database.

The CCOS data are being archived and made available by the CARB. However, much of the data have not undergone a complete quality assurance analysis and, as such, will require that during the SIP project they be analyzed as they are used.

This section provides a brief overview of the CCOS study and its database. The data available during both of the episode periods, both IOP and non-IOP, are described. A few additional data sources will also be used for producing and evaluating modeling inputs. These sources are identified and briefly described in this section as well.

CCOS Field Study

The CCOS was a large-scale field program involving many sponsors and participants with a research budget of over $8 million for the summer 2000 field measurement campaign. In addition, the CARB and local Air Quality Management Districts (AQMDs) provided

substantial in-kind contributions during the measurement campaign. The CARB was responsible on a day-to-day basis for management of the study.

The CCOS field measurement program covered a domain that includes much of northern California, extending north of Redding, and all of central California, including the San Francisco Bay Area and the San Joaquin Valley. A summary report on the CCOS field operations has been completed (DRI, 2001) and is available online: http://www.arb.ca.gov/airways/ccos/docs/ccosv3fdS0.zip. For background information, this section provides a brief overview of the data collected during CCOS. For more details, the reader should consult the summary report.

Study Period

The **primary study period** for CCOS extended from 7/6/2000 to 9/30/2000. During that period, continuous surface and upper-air meteorological measurements and surface air-quality measurements were made for ozone (O_3), nitric oxide (NO), oxides of nitrogen (NO_x), reactive oxidized nitrogen (NO^*_y), nitrogen dioxide (NO_2), peroxyacetylnitrate (PAN), and other peroxyacetylnitrates, particulate nitrate (NO_3^-), formaldehyde (HCHO), and speciated volatile organic compounds (VOC) from automated gas chromatography with ion-trap mass

spectrometer (three research sites). At regular intervals, speciated VOC were also available during the primary study period from PAMS.

During the **intensive operation periods** (IOPs), additional measurements were collected including instrumented aircraft measurements, speciated VOC at more locations, and radiosonde and ozonesonde measurements. During the month of August only, an ozone LIDAR was deployed at Livermore, measuring vertical ozone profiles from 50 m to 2000 m with a 200 m-range resolution.

Routine Data

The routine data available during summer 1999 were also available during summer 2000. The data for 2000 are incorporated into the Central California Air Quality Study CCAQS database.

Field Study Data

The CCOS field measurement program consisted of four categories of surface measurement sites: "supplemental" (S) sites consisting of Type 0, 1, and 2 sites; and "research" (R) sites. The measurements made at each type of supplemental monitoring site are tabulated below. One of the S1 sites was a mobile van operated in the vicinity of Livermore. The carbonyl measurements and the speciated HC measurements at all but the research sites were only collected during the IOPs.

- **Type S0 Sites:**
  - O_3, NO, NO_y
  - wind speed, wind direction
  - temperature, and relative humidity

- **Type S1 Sites:**
  - S0 measurements, plus CO, CO_2, speciated HC, carbonyls

- **Type S2 Sites:**
  - S1 measurements, plus NO_2, PAN

- **Research Sites (3):**
  - S2 measurements, CO, CO_2, NO_y^*, particulate nitrate,
  - light absorption, scattering, actinic flux

Six profilers with RASS were installed and operated during summer 2000 as part of the Central Regional Particulate Air Quality Study (CRPAQS). In addition, nine profilers with RASS and 5 sodars were installed for the CCOS summer 2000 field study. Another sodar was located in the vicinity of the Pittsburgh power plant stacks. During IOPs only, radiosondes and ozonesondes, one in the Sacramento Valley and one in the San Joaquin Valley, were deployed six times per day.

Four instrumented aircraft were used to measure the vertical and horizontal gradients of temperature, humidity, and pollutant concentrations in the study region during CCOS IOPs. These aircraft included a Cessna 172RG and a Cessna 182 operated by University of California, Davis (UCD), and a Cessna 182 and Piper Aztec operated by Sonoma Technology, Inc. (STI). One additional aircraft (Twin Otter), flown by the Tennessee Valley Authority (TVA), made measurements in power plant plumes. The TVA data were collected to evaluate the plume-in-grid parameterizations used in air quality models.

Supplemental Data

A number of supplemental data sources exist that may be useful to this study. For example, an on-road vehicle remote sensing special measurement study was conducted by CARB and coordinated with the CCOS study; the CARB also contracted UC Davis to conduct a vehicle traffic count study; and Districts supplied day-specific plant schedules and pollutant profiles, when available. These data will be used for checking the modeling emissions inventory estimates if they are analyzed and available on a schedule that is consistent with the schedule for this project.

Other data sources were independent of CCOS and mostly the result of routine data collection and analysis efforts. These included synoptic-scale meteorological analysis products and satellite data from multiple platforms and sensors. The meteorological analysis products will be used as inputs to the meteorological model; the satellite data yield products that will provide inputs to both the meteorological model and the photochemical model. The meteorological model requires surface vegetation amounts, and sea surface temperature; the photochemical

model requires total ozone column and surface albedo. These inputs can be derived from satellite data products and/or standard information from the U.S. Geological Survey.

CCAQS Database

According to CARB, all CCOS field observations are currently available from CARB, though not all are in the airways database (http://www.arb.ca.gov/airways/Datamaintenance) and the data formats are not uniform. The level of quality assurance checking conducted for these data is variable; most data have not undergone a complete analysis. To the extent possible, given the time constraints of the project, data will be validated as they are used.

3. MODEL SELECTION

An emissions, meteorological, and photochemical air quality modeling system has been selected that we believe best meets the District's needs in providing high quality modeling databases that can be used for developing the 2004 ozone SIP control plan for the SFBA. This belief is based on the technical features of the selected modeling system and its ability to address the challenges of modeling in the SFBA, the experience and capabilities of the District staff, and the need to maximize the likelihood of a successful model application that achieves the model performance objectives. Specifically, the system we propose comprises the EMS-95 emissions processing model, the RAMS meteorological model, and the CAMx photochemical model.

CHALLENGES OF THE STUDY

There are numerous challenges related to air quality modeling of the Bay Area that will have to be overcome in performing this work effort.

Meteorology: The meteorology of the SFBA and surrounding regions in the CCOS domain is quite complex, and appropriately simulating the effects of micro-climates and flow regimes will be a significant challenge that requires the attention of experts, experienced modelers, and a state-of-science meteorological model:

- Land/sea/bay breezes
- Mountain/valley wind systems in complex terrain
- Role of maritime stratus
- Mesoscale eddies
- Low-level jets

Emissions: Emissions modeling of the Bay Area and central California presents a challenge due to the multitude of diverse sources and the need to remain consistent with the CARB's emissions data and modeling system. Thus, the CARB's emissions modeling system is needed along with full knowledge of how CARB staff generate their emission rate estimates and spatial surrogates:

- On-road mobile sources
- Non-road sources
- Area sources
- Refinery and other industrial sources
- Electric generating sources
- Biogenic emissions
- Quality assurance and quality control (QA/QC)

Photochemical Modeling: The challenges of the meteorological and emissions modeling of the Bay Area are combined with additional chemical and physical challenges in the photochemical modeling. A state-of-science photochemical grid model with the latest model sensitivity analysis capabilities will be needed to address this component, along with the use of:

- Multiscale two-way nested grid resolution (e.g., 1/4/12-km)

- Sufficient vertical resolution
- Current chemical mechanisms (updated CB4, SPARC99)
- Efficient and accurate numerical solvers
- Accurate and mass consistent interface between the meteorological and photochemical grid models
- Probing tools such as Process Analysis, Decoupled Direct Method of sensitivity tracking, and Ozone Source Apportionment Technology

Regulatory Issues: The ultimate objective of the study is to develop a photochemical modeling database that can be used for the year 2004 1-hour ozone control SIP. This SIP must satisfy:
- EPA's SIP guideline documents and requirements including those for photochemical modeling (EPA, 1991; 1996; 1999)
- CARB's guidance documents including those for photochemical modeling (ARB, 1992)
- Continuous contact with the CARB to assure that the modeling meets CARB's approval
- Continuous contact with EPA to assure that the modeling is performed to level that leads to an approvable SIP

Strategic Issues: The modeling and computer systems to be set up to address the 2004 1-hour ozone SIP will be applicable to numerous other air quality issues that will be needed in the future:
- The District will be able to use the system to develop a historical ozone modeling "climatology" and to analyze SFBA impacts on downwind areas due to transport over a wide range of episodes.
- The modeling system will be directly applicable for addressing 8-hour ozone when EPA issues the final 8-hour ozone implementation plan.
- A photochemical model that includes an advanced particulate matter (PM) treatment can be readily adapted to treat fine particulate and visibility issues
- The modeling and computer system will be powerful enough to perform real-time ozone forecasting for the Bay Area.

SELECTED MODELING SYSTEMS

The modeling components selected for the SFBA 2004 SIP revision were specifically identified and requested by the BAAQMD before the study was initiated. All of the models recommended by the District are considered state-of-the-science, and District staff possess a sound experience base for most of the modeling components. All of the selected models have been, or are currently being, used nationally for various ozone, carbon monoxide, and PM SIPs and/or regional regulatory analyses, and thus have been accepted by the EPA and many States for this purpose.

Emissions Model: The processing of episode- and grid-specific emission estimates must use the CARB's emissions data and modeling system, which is based on a California version of the 1995 Emissions Modeling System (EMS-95). Use of any other processing system would result in inconsistencies with ozone SIP modeling in other areas of the CCOS domain (e.g.,

San Joaquin Valley) and could produce conflicting results (e.g., inconsistent conformity budgets). Thus, use of EMS-95 is an essential component of the modeling system.

Meteorological Model: Either the RAMS or MM5 prognostic meteorological models would be the most logical choice for this component of the modeling system. Both models are state-of-science, have a large user community, and are available to all public agencies. We believe that RAMS provides a better treatment of the highly non-hydrostatic processes associated with mesoscale land/sea/lake breeze and planetary boundary layer (PBL) circulations in complex terrain. We have selected RAMS over MM5 because District staff have used this model for several years and so are quite familiar with it, it has demonstrated good performance in the Bay Area, and it provides more flexible grid nesting arrangements (MM5 is limited to a 3:1 ratio when using it in 2-way nested mode).

Photochemical Grid Model: The three logical candidate photochemical grid models for this study include Models-3/CMAQ, CAMx, and UAM-V. The status of UAM-V in terms of public availability is not clear, and its access is strictly guarded. It is also based on legacy (1970-80s) computer code, with very little updating over the past 5-7 years. Thus, UAM-V would not be a good choice. Both CAMx and Models-3/CMAQ are modern codes (1995+) that incorporate state-of-the-science features for all physio-chemical processes. For this study we have selected CAMx over CMAQ because:

1) CAMx can accept meteorological input fields derived from any meteorological model, while CMAQ is limited to the use of MM5;
2) CAMx supports two-way grid nesting at any nesting ratio (e.g., 2:1, 3:1, 4:1), whereas CMAQ supports only one-way nesting at a ratio of 3:1;
3) CAMx has demonstrated good ozone model performance in southern California (Morris et al., 2002), whereas to date only some limited CMAQ modeling for California has been undertaken;
4) CAMx has demonstrated successful application in several ozone SIP modeling studies nationally, whereas CMAQ has not yet been used in an ozone SIP;
5) CAMx supports multi-processing capability to speed execution, which is not in the current version of CMAQ;
6) CAMx supports a full suite of probing tools (DDM, OSAT, and Process Analysis) that may be important in insuring that the model is working correctly, whereas CMAQ just supports Process Analysis;
7) the District has a greater familiarity with CAMx and has used it before; and
8) the project team's familiarity with the model will ensure that a working, fully acceptable modeling system will be developed.

4. METEOROLOGICAL MODELING

The ENVIRON team will use the Regional Atmospheric Modeling System (RAMS) as the prognostic meteorological modeling component of the air quality modeling system to develop the meteorology for the three SFBA ozone episodes (plus initialization days). RAMS has been used for this type of simulation for almost 20 years and the application of RAMS by ATMET personnel (which include the original developers of RAMS) will ensure that the District will attain acceptable meteorological simulations. The District has been using RAMS for many years and are familiar with its application. By necessity, the CAMx air quality modeling domain and grid specifications will be based on CARB's current modeling projection configuration, which is a fairly large regional domain on a Lambert Conformal Conic projection. RAMS operates on a Rotated Polar Stereographic projection; thus, an intermediate processor will be used in the RAMS/CAMx modeling system to perform the necessary manipulations of the RAMS output to properly feed into CAMx.

DESCRIPTION OF RAMS

RAMS has many advantages that make it attractive for these types of simulations:

(a) **Non-hydrostatic formulation:** RAMS uses a compressible, time-split non-hydrostatic equation set in its formulation. The predecessor code to the current RAMS started as a non-hydrostatic model more than 20 years ago. Aside from the model's ability to run on arbitrarily high resolutions, our experience has shown a major benefit to the use of this type of non-hydrostatic code even for larger scale simulations. A compressible non-hydrostatic model will adjust much more quickly to the introduction of observations through the initial conditions, boundary conditions, or the 4-dimensional data assimilation schemes.

(b) **Wide range of physical parameterizations:** RAMS is a very general meteorological simulation system that can be applied to a wide range of atmospheric motions ranging from a hemisphere down to the microscale where boundary layer eddies can be resolved. It contains the physical parameterizations necessary to handle these scales and includes a full suite of radiative, convective, microphysical, vegetative, and soil schemes.

(c) **Flexible domain configurations:** RAMS contains a very flexible two-way interaction grid nesting scheme. Any number of nested grids can be specified at any spatial resolution ratios. The nested grids can either be telescoping or have more than one grid that share the same parent grid. If the application warrants, the nested grids may also move in time. Nested grids may be higher resolution in both the horizontal and vertical and run with a user-specified smaller timestep.

(d) **Sophisticated data analysis scheme:** RAMS uses a hybrid isentropic/terrain-following data analysis scheme (ISAN) to prepare the observations for use as the model initial conditions or in the 4-dimensional data assimilation scheme. This type of coordinate system has been shown by NOAA/FSL and others to generate superior data analyses

when compared to a standard pressure or height analysis. ISAN also allows a wide range of input observation types including rawindsondes, surface observations, towers, wind profilers, buoys, etc.

(e) **Experience with regional scale simulations:** RAMS has been used for more than 10 years to supply meteorological fields to photochemical models. Among numerous other efforts, RAMS was used for the Lake Michigan Ozone Study program and in the OTAG simulation efforts. We have already encountered and solved the interface issues in linking the RAMS fields to photochemical models, including UAM-V and CAMx.

(f) **Ongoing developments:** RAMS continues to be developed and new features are added frequently. Over the past year, a major development has been to add an option for a new vertical coordinate, similar to the "ETA" coordinate. This allows RAMS to simulate arbitrarily steep topography, in addition to doing extremely high resolution runs (grid spacings of 1 m or less) of flow around buildings and other structures.

RAMS has been developed by a number of groups since its inception, including Colorado State University (CSU) and Mission Research Corporation (MRC). With the changes over the past year, the primary focus of development will be at ATMET and Duke University, although CSU and MRC will still be involved. RAMS is a multipurpose, numerical prediction model that simulates atmospheric circulations ranging in scale from an entire hemisphere down to large eddy simulations (LES) of the planetary boundary layer. It is most frequently used to simulate atmospheric phenomena on the mesoscale (horizontal scales from 2 km to 2000 km) for applications ranging from operational weather forecasting to air quality applications to support of basic research. RAMS has often been successfully used with much higher resolutions to simulate boundary layer eddies (10-100 m grid spacing), individual building simulation (1 m grid spacing), and direct wind tunnel simulation (1 cm grid spacing). RAMS' predecessor codes were developed to perform research in modeling physiographically-driven weather systems and simulating convective clouds, mesoscale convective systems, cirrus clouds, and precipitating weather systems in general. RAMS' use has continued to increase to more than 160 current RAMS installations in more than 40 different countries. Although RAMS is supported on all UNIX, Linux, and Windows platforms, because of the exceptional price/performance ratios, we are recently focusing on Linux PCs and PC clusters as our primary computational platform.

The current version of RAMS that is released to the general RAMS user community is version 4.4. We anticipate that version 5.0 will be released in late 2002. Along with an upgrade of the RAMS code structure to more modern and safer FORTRAN 90 constructs, during the time frame of this project, the following features will have been added to the v5.0 RAMS code:

- Generalized observational-nudging 4DDA scheme
- Enhanced analysis nudging options
- Antecedent precipitation index scheme for soil moisture initialization
- Several diabatic initialization options
- Use of NDVI datasets to define vegetation characteristics

RAMS CONFIGURATION

We will employ RAMS v4.4 as the control run for the three episode cases in 1999 and 2000. In each case, the actual RAMS runs will begin 72 hours before the first episode day to provide the spinup days for CAMx. While various aspects of the model configuration may change upon our review of the specific meteorology, observation availability, and air quality characteristics, the following discussion approximates what we anticipate the RAMS configuration will look like.

Modeling Grids

For the simulations of the three episodes, the RAMS grids will be configured similarly to previous simulations performed by CARB with MM5. We will primarily use a three grid nested structure with the finest grid at 4 km resolution. However, sensitivity tests using a fourth grid with a 1 km resolution over the Bay Area will also be tested to determine the improvements with even higher resolution of the coastlines and topography. Surrounding the finer grids will be a 12-km nest, which in turn will be nested within a 48 km grid to resolve the large scale forcing. Our experience has shown that the meteorological results are greatly enhanced if a significant portion of the synoptic scale is included in the simulation domain, rather than just forced in through the boundary conditions or the four-dimensional data assimilation scheme.

For the vertical structure, RAMS will be configured to run all grids with 41 coordinate levels with the lowest wind and temperature level at about 10 m AGL then smoothly stretching to a maximum of about 1000 m grid spacing. The top of the model will be placed at about 20 km MSL to ensure that the various synoptic scale features such as the sub-tropical jet stream (which is located about tropopause level) are adequately resolved in the simulation domain. Although the upper level jets are not directly important in the low-level transport of ozone and its precursors, the jets do affect the low-level pressure patterns which control the low-level winds.

Care will be taken to closely coordinate the RAMS and CAMx grid resolutions and domain coverages to minimize the impact of interpolation errors and better ensure mass-consistency in the transfer of the meteorological fields from the meteorological model to the photochemical model.

The 4 km CCOS air quality modeling domain is shown in Figure 4-1. While the CARB is performing their modeling on this entire domain at 4 km spacing, we will configure the RAMS grids to cover the focus region (SFBA) at 4 km resolution, with 2 coarser nested grids surrounding it. Figure 4-2 and Table 4-1 depict a likely configuration for the grids. The coarsest grid is 48 km resolution, with grid 2 covering the entire state of California at 12 km spacing. The figure and table also depict an optional grid 4 with ultra-high resolution, which will be employed for sensitivity simulations. Two-way nesting communication will be used for all grids. Figure 4-3 shows the expected vertical layer structure for all RAMS grids; we have defined the layer structure in the lowest 1500 m to be identical to that proposed by CARB for their CCOS modeling efforts.
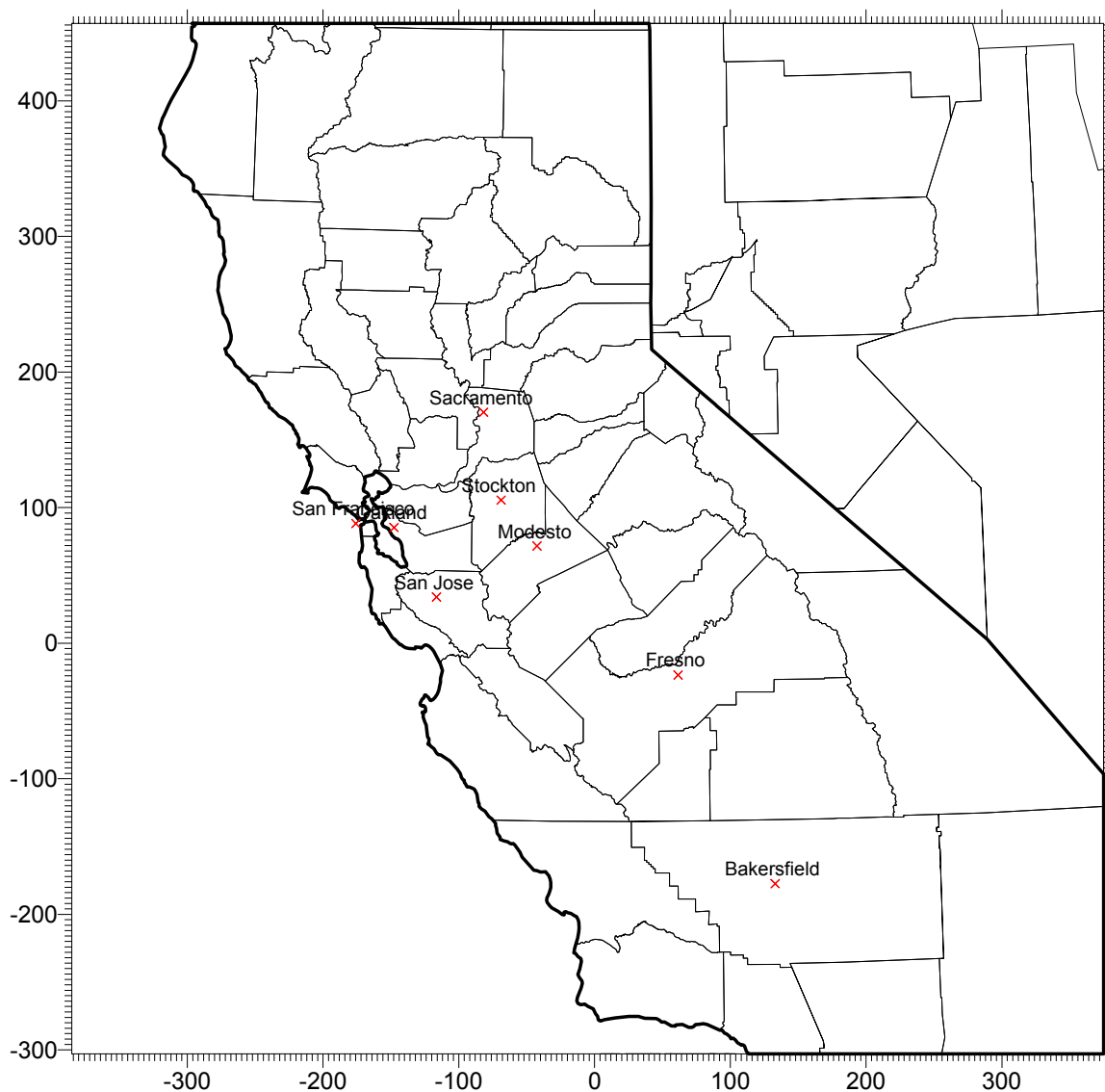
CCAQS 4km Grid
LCP Center: (120.5W, 37N)
Standard Parallels: 30N, 60N
SW Corner: (-385.13, -302.91)
NX x NY: 190 x 190

Figure 4-1. The coverage of the CARB/CCOS air quality modeling domain. Grid spacing over the entire region is 4 km. Map projection is Lambert Conformal.
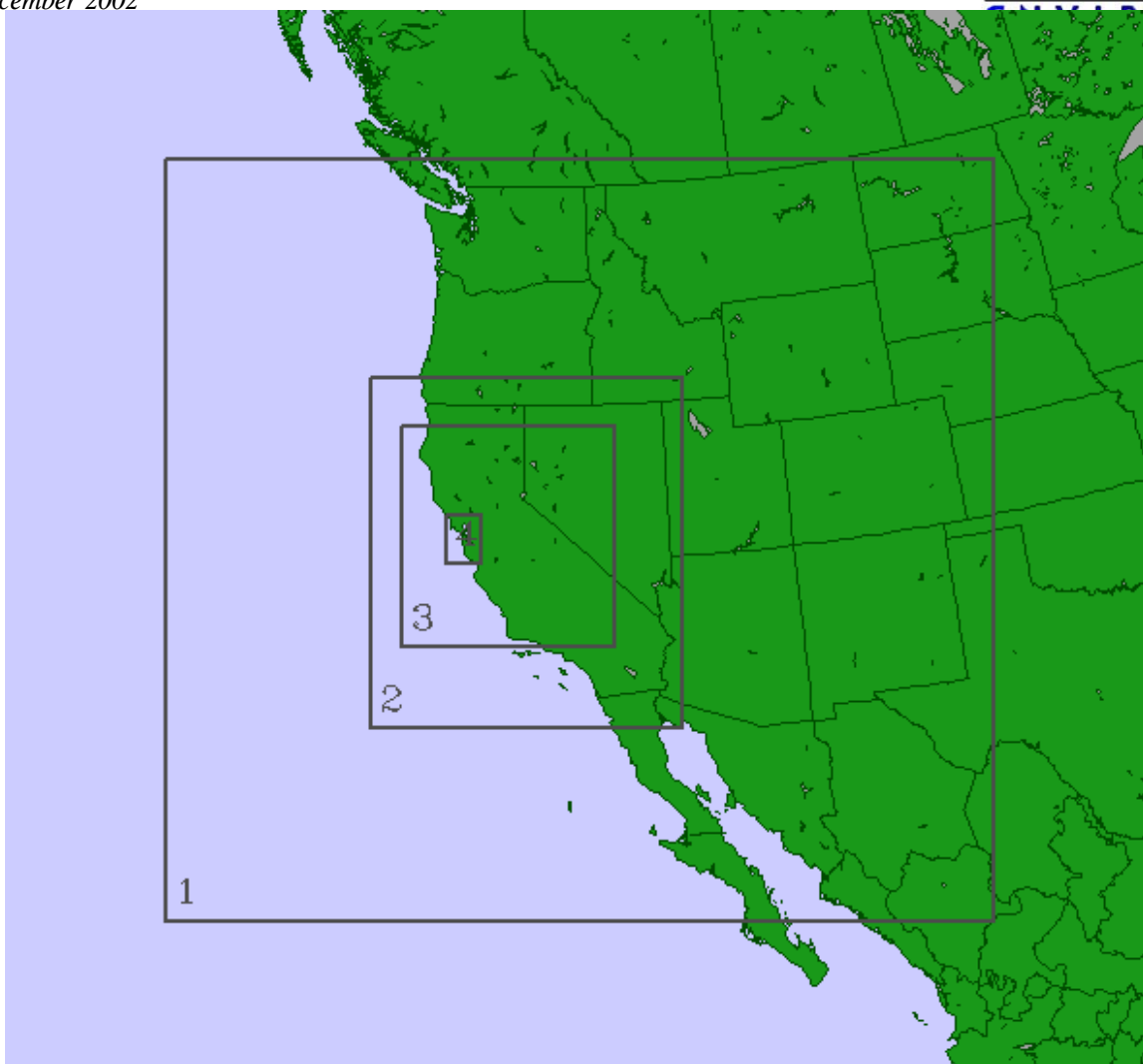
Figure 4-2(a). Example of the likely configuration for the RAMS rotated polar stereographic modeling grid, which will employ a system of up to four nested grids with successively finer resolution.
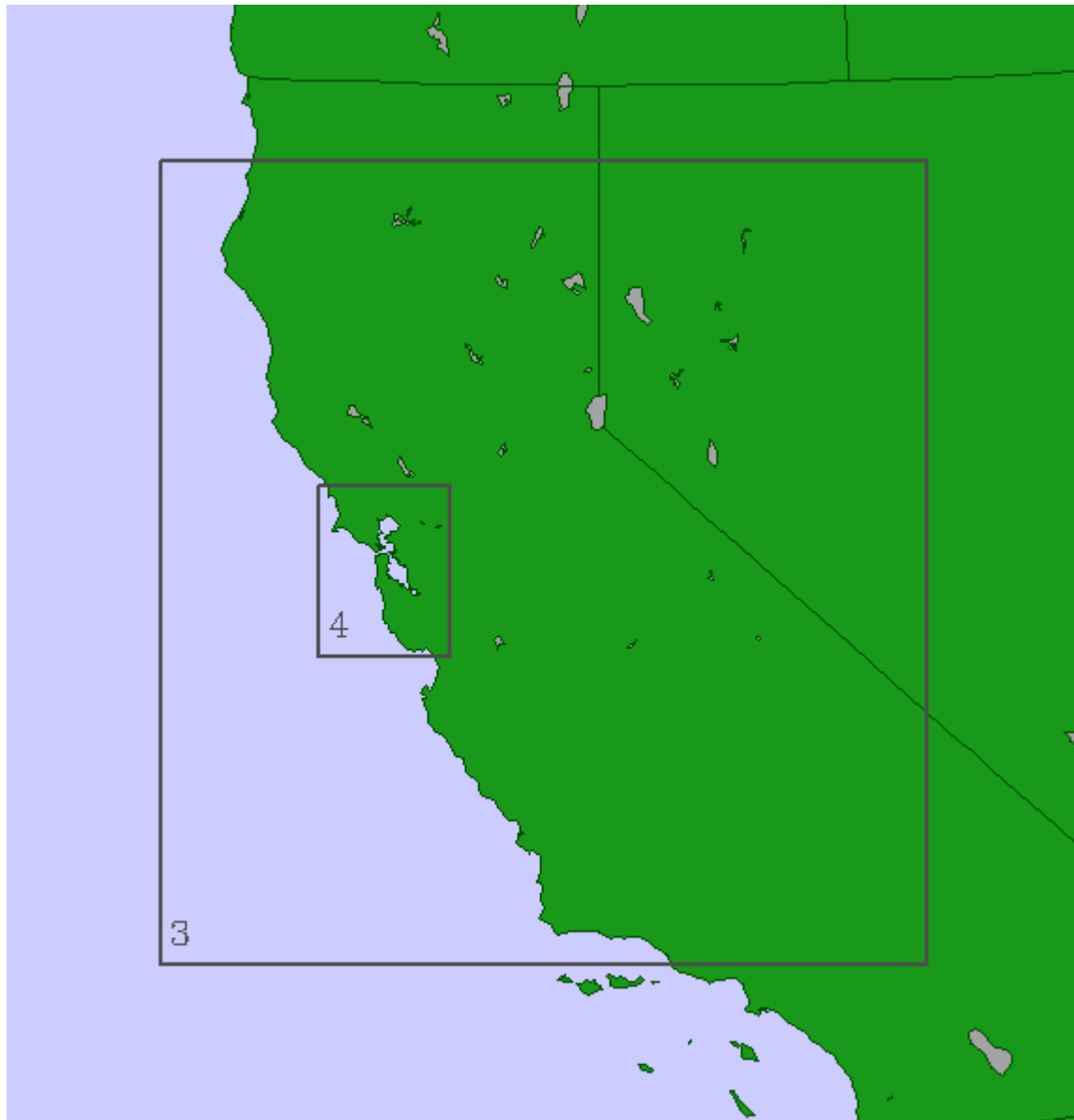
Figure 4-2(b). Blowup view from Figure 4-2(a) above, showing the innermost two RAMS grids.
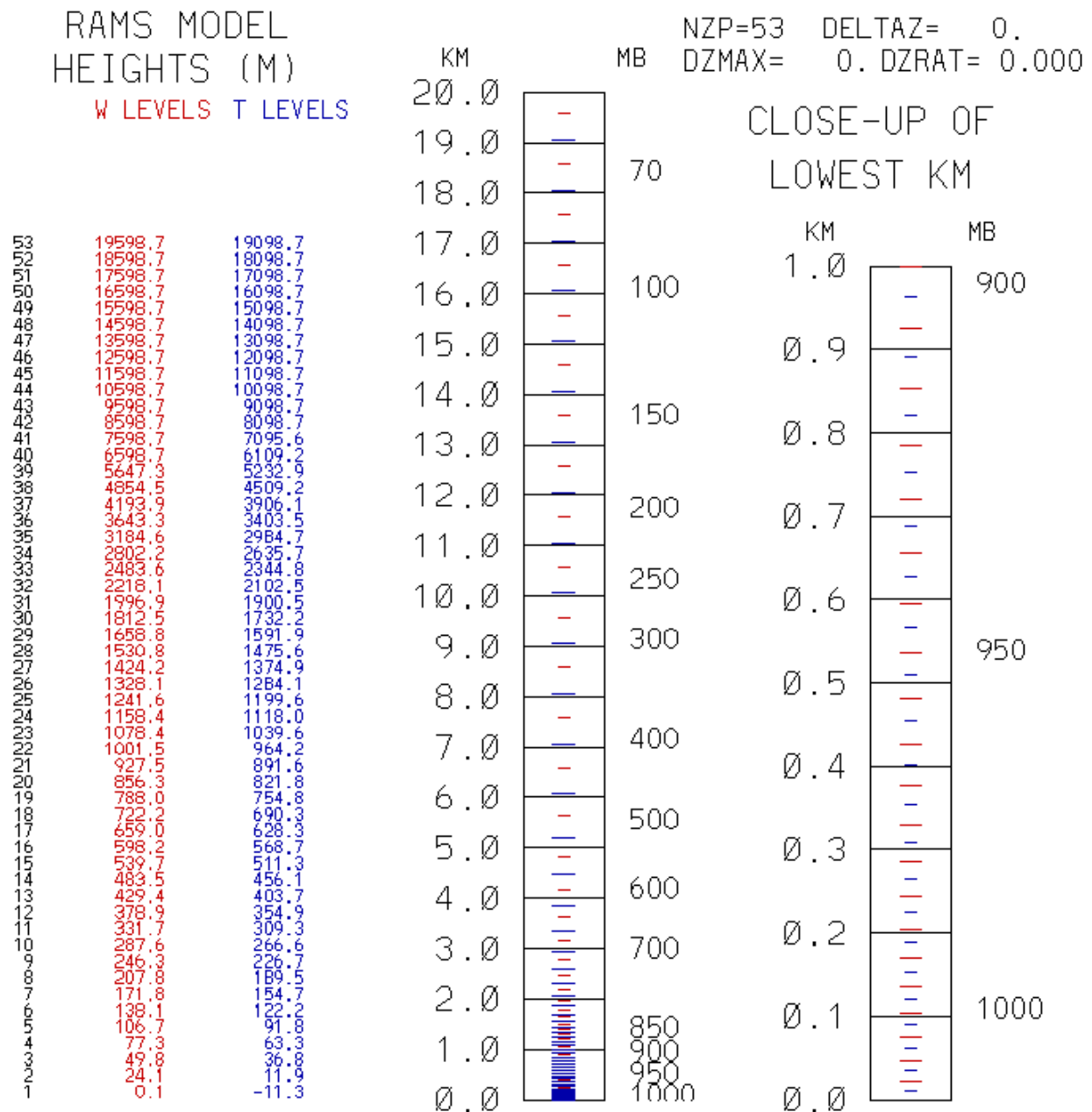
Figure 4-3. Expected vertical structure of RAMS grids.

ENVIRON

Table 4-1. Grid parameters for each of the nested domains shown in Figure 4-2.

Grid	# of X points	# of Y Points	Vertical Levels	Δx (km)	Δy (km)	Δz (m) (Lowest)
1	63	58	53	48	48	24
2	94	106	53	12	12	24
3	191	200	53	4	4	24
4	130	170	53	1	1	24

Input Data

The input meteorological data for the episodes will be derived from standard datasets along with the available special observations from the CCOS. The meteorological input data to the meteorological models can be grouped into three categories:

1) **Large scale gridded analyses:** Global analyses of meteorology are available from the National Centers for Environmental Prediction (NCEP). We will use the NCEP/NCAR Reanalysis data. The parameters of wind, temperature, and humidity are analyzed on pressure levels (20 levels extending from 1000 mb up to 10 mb) on a 2.5 degree latitude-longitude grid. These data are archived every 6 hours and serve as a first guess field for the data analysis. We will access this data from the National Center for Atmospheric Research (NCAR).

2) **Standard NWS observations:** The rawinsondes and surface observations reported by the NWS and other national meteorological centers are also archived at NCAR. The rawinsondes are reported every 6 hours and the surface observations are archived every three hours. These data will be accessed for the 6 day period.

3) **Special observations:** Special observations taken during the summer of 2000 from the CCOS monitoring sites will be included in the data analyses and FDDA. These observations included surface observations, wind profilers, rawinsondes, etc. Furthermore, team member Dr. Robert Bornstein has been deeply involved in the development of the Bay Area Mesonet Initiative (BAMI), a program that automates the collection, quality assurance, and consistency of local meteorological data from several networks for public dissemination. To the extent that BAMI data are available during the summer of 2000, and that they do not overlap CCOS databases, we will investigate and consider all observational data for inclusion into the simulations for the three episodes.

The Bay Area Mesoscale Initiative (BAMI) is a consortium between San Jose State University (SJSU), the Naval Postgraduate School (NPS), and the National Weather Service Office in Monterey (NWS) that seeks to enhance the understanding of regional meteorology through establishment of a real-time web-based distribution of meteorological data from mesoscale observational networks. Many agencies already gather such data for different purposes, and

BAMI aims to gather and distribute those data. The Cooperative Program for Operational Meteorology, Education, and Training (COMET) in June 1999 thus awarded the consortium funding to use BAMI mesonet data in Limited Area Prediction System (LAPS) for local mesoscale modeling and forecasting. The project has unified the data formats of each constituent network, eliminating non-meteorological information, and provided data in a uniform format for easy processing. The NPS retrieves data from various institutions and SJSU has developed and maintained the BAMI web site (http://meso.met.sjsu.edu/bami), but data are distributed through the Local Data Acquisition and Dissemination (LDAD) system of the Western Region NWS. The LDAD system allows Weather Forecast Offices to collect, check and share local mesonet data, not only within the NWS, but also with local emergency management agencies. BAMI data are also sent to MesoWest, a BAMI-type project for Utah and the western U.S. BAMI data are ingested in the MesoWest data flow to the LDAD system, which makes them available to the entire country in several ways. The simplest is the Forecast Research Laboratory display at http://www-frd.fsl.noaa.gov/mesonet/.

Upon acquisition of the observational data, all NWS, CCOS, and possibly BAMI observational data will be processed with our quality control algorithms. We have developed a QC package which consists of three separate schemes: 1) internal consistency checks, 2) "buddy" checks, and 3) "first-guess field" checks.

The internal consistency checks consist of basic sanity and range checking of the observational data along with the physical constraints of hydrostatic balance. Rawindsondes are also checked for lapse rate and wind shear realism. The buddy checks will compare a station's value with that of its neighboring stations. The checks versus the first-guess fields will compare an observation against the large scale gridded pressure data analyses. At any of these three stages, observational data values can be flagged as missing, bad, suspect, or corrected.

After the input meteorological observational data has been quality-controlled, it will be combined with the large-scale gridded analyses to produce a complete data analysis for RAMS initial conditions and the 4-dimensional data assimilation scheme. RAMS/ISAN (Isentropic Analysis package) will be used for the analysis. ISAN is a hybrid isentropic/terrain-following height coordinate scheme which uses a Barnes-type objective analysis algorithm.

Other types of input data which describe the surface characteristics are also necessary for the execution of RAMS. We already possess archives of high-resolution topography, land use, and NDVI for the entire domain. These datasets are global and have about a 1 km resolution.

RAMS Physics and FDDA Configuration

We expect that RAMS will be configured with the following physical and numerical options for the CCOS runs:

- Mellor-Yamada type diffusion coefficients with prognostic turbulent kinetic energy
- Long and short wave radiative parameterizations
- Prognostic soil temperature and moisture model
- Prognostic vegetation parameterization

- Explicit and parameterized precipitation
- Four-dimensional data assimilation (analysis and observational nudging)

The four-dimensional data assimilation (FDDA) scheme which has been used in the past by RAMS for these types of simulations has been termed in the meteorological literature as "analysis nudging". However, in certain circumstances, "observational nudging" has some advantages. With the new observational nudging scheme that has been implemented in RAMS, we will have the ability to exercise and test the sensitivity to both types of FDDA schemes.

OUTPUT AND EVALUATION

RAMS will be set to output the simulation results every hour. A complete set of fields will be output for all model grids, including u, v, w wind components, temperature, pressure, cloud variables, precipitation, and eddy diffusion coefficients (or turbulent kinetic energy). The RAMS output files will be converted to CAMx-ready files for use in the photochemical model.

The RAMS Evaluation and Visualization Utilities (REVU) package will be supplied as part of the overall system. For graphical depictions of the meteorological fields, REVU uses NCAR Graphics to generate plots which can then be converted to various other formats such as Postscript, GIF, etc. In addition, we will also install RINGI (RAMS Interactive NCAR Graphics Interface) as part of the system. RINGI is a Graphical User Interface based on TCL/TK, built on top of the REVU package, which provides a convenient, interactive way to look at the raw RAMS output files with the standard REVU NCAR Graphics plots. REVU also has the ability to convert RAMS output files to Vis5D or GrADS format if desired.

For the model's statistical performance, we will use a recently developed, generalized statistical package designed by MRC and ATMET which we termed REVU-GS. Based on the RAMS/REVU code, REVU-GS is compatible with the RAMS input observation files and the raw RAMS output files. The package can produce graphics and tables of various statistical measures for any number of times and levels during a simulation, enabling both spatial and temporal verification. Among the statistical parameters that can be calculated by the REVU-GS package are:

- Root mean square error (RMSE)
- Mean absolute error
- Relative error
- Bias
- Root mean square vector error (RMSVE)
- Correlation coefficient

REVU_GS also has the ability to subsample the observations in order to compute the statistics over different regions of the domain. The sampling can be done by user-defined latitude-longitude bounds, on a nested grid, or by physically-based subdomains (e.g., coastal versus inland versus mountain stations).

Along with the statistical performance measures, we will also phenomenologically evaluate the meteorological results for the presence of expected circulations over the northern and central California region. This type of evaluation is dependent on the availability of various types of observed data for these events. For example, we have access to sources of hourly precipitation data and national NEXRAD mosaics for these events. There also are sources of at least 12 hourly visible satellite images for the year of 2000. We will continue to look for higher time-resolution visible images that may be available to better evaluate how well RAMS was able to handle fog and low clouds. Although there is little quantitative information, we can assess how well the areal coverage was represented in the model. We will also assess the availability of other specialized satellite products such as IR and skin temperature images, although the presence of fog and low clouds can make these products difficult to interpret.

Other types of data would also come in useful for evaluation, including sodar and RASS data to estimate mixing heights and wind profiler information to investigate the vertical structure of the various circulation features. Once we have deduced the total availability of observational data, we will have a better idea of the process to make the phenomenological evaluation, along with investigating how well the RAMS simulations performed in capturing the typical circulation features of the low-level jets, sea breezes, slope flows, and the Schultz and Fresno eddies.

If time permits in the project, other tools will be used for the evaluation. A companion model to RAMS called HYPACT (HYbrid Particle And Concentration Prediction model) is a fully three-dimension transport and dispersion code which can operate in several modes, including standard Lagrangian particle dispersion mode, trajectory mode, and Eulerian mode. Although chemical transformations are not included in the current version of HYPACT, the use of trajectory analysis could come in useful is helping to asses the realism of the various circulation features.

Sensitivity Simulations

Numerous sensitivity runs of each case will be performed to demonstrate model integrity, sensitivity to resolution, physics, 4DDA, use of special observations, etc. The exact structure of these sensitivity simulations will be discussed with District and the MAC. We will begin with a three-grid control case simulation using RAMS v4.4. These results will then be compared to the new v5.0. We expect the v5.0 results to be similar or better than v4.4, as completed tests have shown in other situations. Additional tests will then be made with v5.0, testing the effect of resolution on the Bay Area with the added 1 km grid 4, along with testing of the new FDDA schemes.

Additional RAMS simulations will likely be required after initial air quality simulation results have been reviewed. The results will be reviewed by the Internal Technical Review Team and the District.

DELIVERABLES

The deliverables for this task are as follows:

- A report on the development, configuration, and evaluations of the meteorological simulations of the three episodes. The report will be posted to the Contractor-maintained web site.

- The meteorological modeling system comprised of all the software components mentioned above. The software components are:
  a) RAMS, including the ISAN component
  b) Quality control package
  c) REVU and REVU-GS
  d) RINGI
  e) RAMS Grid Configurator
  f) Any supporting components for data conversion
  g) GrADS
  h) Vis5D
  i) NCAR Graphics
  j) RAMS to CAMx converter

5. EMISSIONS MODELING

In order to effectively conduct air quality modeling for the Bay Area SIP revision, it is necessary to develop temporally and spatially resolved emission estimates that are suitable for input to the photochemical model. Emissions are broadly categorized into major stationary or point sources, area sources, on-road mobile sources, non-road mobile sources, and biogenics. In addition, there are many subcategories that comprise the non-road mobile, point and area sources. In the following section, we describe the emissions model that will be used in this study. We also describe where the emissions data will be obtained and how they will be used to develop base case and future year emissions estimates.

EMS-95

In order to remain compatible with on-going activities at the CARB, we will use the 1995 Emissions Modeling System, or EMS-95 (Dickson and Oliver, 1991; Dickson et al., 1992; Bruckman and Oliver, 1993; Wilkinson et al., 1994; Janssen, 1998), to prepare the spatially, temporally, and chemically resolved emissions estimates of total organic gases (TOG), oxides of nitrogen (NO_X), and carbon monoxide (CO) for the point and area sources. EMS-95 is the emissions modeling system that is currently used by the CARB. Though EMS-95 is capable of preparing biogenics and on-road mobile source emissions estimates, the CARB uses separate systems to prepare these estimates.

For biogenics, the CARB uses the Biogenic Emission Inventory Geographic Information System, or BEIGIS (CARB, 2001). For on-road mobile sources, the CARB is in the process of developing new, statewide emissions estimates under a GIS framework. For both on-road mobile sources and biogenics, we will use the spatially and temporally resolved estimates of TOG, NO_X, and CO prepared by the CARB for use in the current study. However, we will use EMS-95 to chemically speciate the TOG component of the biogenics and on-road mobile sources into the individual organic species needed by the air quality model's chemical mechanism. Further, we will use EMS-95 to reformat the emissions estimates for input to CAMx.

To expedite the application of EMS-95 for the current study, the CARB will provide a copy of its working version of EMS-95 and BEIGIS on CD-ROM to the project team. This will help ensure that we remain compatible with on-going CARB activities.

Emission Estimates from CARB for the CCOS Period

The CARB will provide emissions estimates for the entire CCOS domain, shown in the previous section as Figure 4-1. The major stationary (point) source inventory for the study domain will contain actual stack coordinates and will include year 2000 ozone season day estimates of TOG, NO_X, and CO for each process of the inventoried facilities. The inventory will include the required elements in the format described in Appendix A (Section A-1). The area source inventory will include year 2000, county-wide ozone season day estimates of TOG, NO_X, and CO for each area source category. The inventory will include the required

elements in the format described in Appendix A (Section A-2). EMS-95 will be used to process the major stationary source and area source inventories into gridded, speciated, hourly emissions estimates suitable for input to CAMx. If available, the CARB will provide day-specific emissions estimates for certain stationary and area sources directly in the pertinent data sets, which are described in Appendixes A-1 and A-2.

The CARB will provide gridded, hourly biogenic emissions of isoprene, monoterpenes, and other volatile organic compounds (OVOC) at a minimum, for the CCOS domain for each day of the episodes.   The CARB will not provide estimates of biogenic nitric oxide (BNO). The inventory will include the required information as described in Appendix A (Section A-3).

Finally, the CARB will provide gridded, hourly on-road mobile source emissions of TOG, CO, and NO_X for the CCOS domain for each day of the episodes. The inventory will include the required information as described in Appendix A (Section A-4). EMS-95 routines will be used to speciate the biogenic and on-road mobile source inventory.

ON-ROAD EMISSIONS FROM EMFAC

EMFAC2001 version 2.08 is the current model used by the CARB to estimate new on-road mobile source emissions factors for California (CARB, 2002). The EMFAC2001 model supercedes EMFAC2000, which was released in November 2000. The CARB will use DTIM4 to develop gridded, hourly, day-specific emissions estimates of TOG, NO_X, and CO for the episodes to be modeled for the Bay Area SIP revision.

Much of the initial work to prepare the base case CCOS on-road mobile source inventory is being done by study team member Alpine Geophysics, LLC (AG). AG has developed an Integrated Transportation Network (ITN) for the entire state of California. The ITN is a combination of twenty-three individual networks from seven transportation planning agencies and CalTrans. For counties that are not represented by a local transportation agency, AG used the CalTrans statewide transportation network. However, unlike other transportation planning agencies, the CalTrans developed estimates of network travel only for personal travel and not commercial travel. Therefore, AG developed estimates of commercial travel for the CalTrans network. Because the individual networks had VMT and trip end data from base years spanning 1995 through 2000, AG had to grow the VMT and trip ends to a common base year – in this case, 2000. Researchers at the University of California at Davis allocated the link-based VMT, which was originally allocated to one or more time periods (i.e., AM peak, PM peak, midday peak, off peak, or daily), to twenty-four hourly bins. AG then allocated the intrazonal VMT and trip ends, in some cases by trip type, also to twenty-four hourly bins. Finally, AG is using EMFAC v.2.08 and DTIM4 to estimate gridded, hourly on-road mobile source emissions estimates for the CCOS domain for the June and August 2000 episodes. At this time, the CARB will be responsible for using the ITN to develop future year emissions estimates for on-road mobile sources.

In all cases, AG used the data provided by the local transportation planning agencies for their specific networks (over other data that were provided by the CARB and CalTrans) to develop the ITN and the resulting CCOS on-road mobile source emissions estimates. When complete

documentation becomes available on the development and use of the ITN, these data will be provided to the study sponsors.

Because the CARB plans to modify EMFAC2001 to support changes to heavy-duty diesel emissions factors, among others, it will be necessary to track these revisions and their potential impact on the on-road mobile source inventory. If the CARB indeed does revise EMFAC2001, the project team will decide with the District how best to proceed with integrating these changes into the air quality modeling inventory.

BIOGENICS

As previously stated, the CARB will use BEIGIS to provide gridded, hourly day-specific biogenic emissions estimates of isoprene, monoterpenes, and OVOC for the episodes to be modeled for the Bay Area SIP revision. It should be noted that OVOC emissions are simply a multiple, approximately 20% to 30%, of the monoterpene emissions estimate. However, the CARB will not provide estimates of BNO. Because biogenic NO has been shown to impact the efficacy of anthropogenic emissions control strategies (e.g. Wilkinson and Russell, 2002), the project team will estimate gridded, hourly biogenic NO emissions using the BEIS3 v.0.9 prototype (Pierce, 2001a). Though other biogenic emissions models are available for use in this study (e.g. BEIS2, BIOME, GLOBEIS), the BEIS3 prototype contains a biogenic emissions landuse database (BELD3) that is more representative of land use conditions of the target episodes (Pierce, et al., 1998). Further, the BELD3 is representative of 230 plant communities and species at a one kilometer by one kilometer resolution over North America. Though the BELD3 is available for use in the study, the project team will evaluate the feasibility of adapting the BEIGIS land use/cover data base for use in estimating BNO, given that the BEIGIS database is specific to California. If the project team determines that it is feasible to adapt the BEIGIS data base into BEIS3 and the project sponsors concur, we will use the BEIGIS data to estimate BNO.

SHIPPING EMISSIONS

NO_x and VOC emissions estimates from oceangoing vessels are substantially underestimated in existing emissions inventories (Corbett et al., 1999; Corbett and Fischbeck, 1997). In order to correct this suspected deficiency, the project team will estimate day-specific NO_x and VOC emissions for oceangoing and San Francisco Bay commercial marine traffic. This effort will require that the study team collect data related to transit activity (i.e., hotelling, maneuvering, cruising, berthing, and auxiliary generator use), time in activity mode, fuel type consumed (e.g., bunker fuel oil, marine diesel), ship type, and engine type. The primary source of this data will likely be the three primary port authorities in the San Francisco Bay Area: The Ports of San Francisco, Oakland, and Redwood City; and the San Francisco Marine Exchange. Navigable waterways data, for use in spatially allocating marine vessel emissions, will be taken from US Department of Transportation (BTS, 2001). The methods to estimate NO_x and VOC shipping emissions will be drawn from Shah (2001) and Trozzi and Vaccaro (1998). Note that only emissions from marine engine exhaust will be estimated, and not from other sources such as ballasting, loading/unloading, or transport of petroleum products.

REFINERY EMISSIONS

Based on work done by the BAAQMD, refinery emissions were increased, specifically from flaring operations, by a factor of one hundred in the recent BAAQMD and ARB emissions inventories (from 0.1 tons per day [tpd] NO_X to 13 tpd NO_X). There is evidence to suggest that other refinery related emissions are also underestimated (e.g. upset events, pressure relief valves). In an effort to better characterize emissions from refinery operations, the BAAQMD has undertaken an effort to develop day-specific emissions estimates for refinery operations in the district. These day-specific emissions estimates will be used in-lieu of the current emissions estimates that reside in the BAAQMD and ARB emissions inventories for the base case air quality modeling. For the future year air quality modeling, the ARB will develop projected emissions estimates for refinery operations.

OTHER EMISSIONS DATA

Other emissions-related data are required in order to prepare emissions estimates that are suitable for input to CAMx. These data include the following:

- Area source spatial surrogates;
- Cross references between area source categories and their spatial surrogates;
- Chemical mechanism-specific hydrocarbon speciation profiles; and
- Cross references between source categories and their hydrocarbon speciation profiles.

The area source spatial surrogates are used to spatially allocate the county-wide area source emissions estimates to individual grid cells. The CARB has developed four kilometer resolved are source spatial surrogates based on census and other data that are representative of conditions prior to 2000. The CARB will use them to spatially allocate the county-wide area source emissions estimates. It is unclear, however, whether the CARB will update the four kilometer area source spatial surrogates based on data such as the 2000 census. Further, the CARB has a contractor currently preparing one kilometer resolved surrogates. If the project sponsors desire to use the finer resolved area source spatial surrogates, the CARB will supply these data for use in the current study. The ENVIRON team has no expectation to develop new surrogates. The CARB will supply these data per the file format described in Appendix A (Section A-7).

The area source spatial surrogates cross reference data maps each area source category to a specific spatial surrogate. The CARB will supply these data per the file format described in Appendix A (Section A-8).

The chemical mechanism hydrocarbon speciation profiles are used to split the TOG emissions estimates into the individual hydrocarbon components that are modeled within the chemistry processes of CAMx. The CARB will provide hydrocarbon speciation profiles for both the CB-IV and SPARC99 chemical mechanisms per the file format described in Appendix A (Section A-10).

The hydrocarbon speciation profile cross reference data maps an area source category or source classification code (SCC) to a chemical mechanism hydrocarbon speciation profile. The CARB will provide this cross reference per the file format described in Appendix A (Section A-11).

Of note, by definition in 40 CFR § 51.10 VOC does not include methane and ethane. However, the chemical mechanism in CAMx does not utilize the methane and ethane components of the emissions; hence, it is necessary to remove methane and ethane from the emissions stream before input to CAMx.

NEVADA EMISSIONS

The ENVIRON team is not aware of efforts by CARB to develop emissions estimates for the portion of the modeling domain in Nevada. It is possible that emissions data from WRAP can be used to address this area. If WRAP is unable to supply emissions data for Nevada, we can use emissions data from EPA. If the project sponsors direct the study team to prepare emissions estimates for Nevada, we will collect the pertinent stationary source, area source, on-road mobile source, non-road mobile source, and biogenics emissions data in order to run EMS-95 for Nevada.

EXTENDING THE SOUTHERN AIR QUALITY MODELING BOUNDARY

If the project sponsors desire to extend the southern air quality modeling boundary further south and CARB is unable to supply adequate emissions estimates for the extended domain, the ENVIRON team will seek to acquire appropriate emissions data to cover the extended domain. The sources of such emissions data include the WRAP and the EPA. Such emissions data will have to be adapted for use in the modeling system.

NEW EPISODE

The project sponsors have expressed a desire to model a July 1999 episode. Because the CARB will not develop emissions estimates for non-CCOS episodes, the project team will have to develop them. The following text outlines the procedures that will be followed in order to develop air quality model ready emissions estimates for the four kilometer-resolved grid for a non-CCOS episode. As stated previously, the study team will require the version of EMS-95, BEIGIS, and supporting data sets that ARB uses so that we can develop the new emissions estimates in a manner consistent with the procedures that the CARB uses to develop the CCOS emissions estimates.

Emissions inventory development for photochemical modeling must address several source categories including (a) stationary point sources, (b) area sources, (c) on-road mobile sources, (d) non-road mobile sources, and (e) biogenic sources. These estimates must be developed for the base year (i.e., the historical year when the ozone episode actually occurred), the future baseline year and the various emissions control scenarios associated with one or more future

base years. Development of base year and projection emission inventories for each of these source categories requires a separate modeling approach as described briefly below.

Development of Stationary Point Source Emissions. There are a large number of electric utility and industrial point sources distributed in a highly non-uniform manner throughout the study domain. Emissions data for these numerous large point sources are typically supplied on a facility level basis. These emissions estimates can be taken from an existing inventory system (e.g., CEIDARS) or from an existing regional inventory (e.g., NEI, WRAP), and then adjusted to the selected base and projection year using a variety of techniques. Annual average or seasonally-adjusted emissions rates of NO_X and other ozone precursors are available for most major point sources from the CEIDARS and NEI inventories and these can be incorporated directly into EMS-95. In addition to emissions flow rates for each pertinent stack or cluster of stacks at a facility, point source inventorying also requires a variety of so-called 'stack parameters', including:

> Data to locate the point source on the modeling grid;

> Operational data to determine the temporal profile of the emissions; and

> Stack configurations to characterize the release point of the emissions.

To screen for simple, but potentially serious inventory errors, AG has developed a number of customized QA reports and programs within the EMS-95 to examine the accuracy of the point source emissions. These reports will be added to the version of EMS-95 that the CARB will make available to the project team for use in this study.

Development of Area and Non-Road Mobile Source Emissions. Area and non-road mobile source emissions are developed from county level emissions totals. The project team will use area source and non-road source emissions data retrieved from the CARB for California and the EPA for Nevada. Treatment of these source categories also requires the application of temporal adjustment factors and spatial allocation factors to disaggregate the county level emissions into an hourly, gridded format. Though generalized temporal and spatial factors will be used to spatially allocate the Nevada emissions, the project team will use the area source spatial surrogates that have been developed by the CARB to spatially allocate the California emissions estimates.

If the CEIDARS/NEI/WRAP base emissions data are within two years of the date of the non-CCOS episode (i.e., 1998 or 1999), we recommend that no backcasting/forecasting be applied to the base year emissions data unless there exists a known major impact to the emissions estimates. It has been our experience that such efforts to backcast/forecast over such a short time frame result in only small changes to the overall inventory, which result in little or no impact to the air quality model results. If backcasting/forecasting is required, then appropriate emissions growth data will be obtained and applied to the base year emissions data representative with

Development of On-Road Motor Vehicle Emissions. On-road mobile source emissions are calculated by combining the estimates of VMT (vehicle miles traveled) by emissions factors specific to vehicle type, operating speed, and ambient temperature. These emission estimates

can be calculated based on county-level estimates of VMT by vehicle type and then distributed along the grid based on a digitized roadway network. If local traffic modeling is available to provide link specific traffic volumes and speed profiles, the emissions can also be calculated based on these specific roadway links. Often the two approaches are combined to provide high-resolution link-based emissions for an urban core, and more generalized estimates for the outlying regions.

For the current study, we will use the ITN to develop on-road mobile source emissions estimates for the non-CCOS episode. We will backcast/forecast the VMT and trip ends for the 2000 ITN to the appropriate episode year. The project team will use EMFAC v.2.2 to also generate new emissions factors for use with the ITN data. Finally, we will use DTIM4 coupled with the salient meteorology to estimate gridded, hourly on-road mobile source emissions estimates for the non-CCOS episode.

Development of Biogenic Emissions. Currently, biogenic emissions are calculated based on day-specific meteorology supplied by the prognostic meteorological model (RAMS or MM5), and gridded land use/land cover data including vegetation types and coverages. We will use the CARB's BEIGIS model and supporting data bases to estimate gridded, hourly biogenic emissions estimates of isoprene, monoterpenes, and OVOCs for the non-CCOS episode. Further, we will follow the same procedure that will be established to estimate BNO for the CCOS episodes to estimate BNO for the non-CCOS episode.

Development of Future Year Baseline and Control Strategy Inventories. Construction of the future year emissions inventories is typically based on a combination of existing data sets and new material developed specifically for the study domain. Projection techniques for developing future-year planning inventories include one or more of the following methodologies:

> Application of growth and control factors developed for regional modeling exercises, such as the CCOS;

> Development of growth factors using EPA approved models such as EGAS;

> Incorporation of growth and control estimates developed by local agencies;

> Incorporation of growth and control estimates developed by local industry; and

> Control factors developed by the CARB in regulatory actions such as the local AQMD SIPs.

Though we know that dramatic changes to land use and land cover have occurred in California that can impact biogenic emissions estimates, on time scales less than a decade, these are difficult to know exactly. Therefore, biogenic emissions will be assumed to remain constant from base year to future year.

Weekend Effects. The July 1999 episode falls on a weekend. There is some concern that the use of an average weekday inventory in the air quality modeling effort has the potential to skew the air quality predictions for an episodic and typical weekend day. The ARB is

currently expending resources to determine the differences in a weekend day and weekday emissions inventory in the South Coast Air Basin and how those differences impact air quality model predictions. We are currently seeking additional information from the ARB in regards to this issue. Specifically, we are attempting to determine if the data that ARB has developed (is developing) can be generalized for use in all of California.

FUTURE YEAR PROJECTION METHODS

Figure 5-1 shows an overview of how air quality modeling emissions inventories are created. In the first step, the base case emissions estimates and data are used to create a historical episode, the "base case" air quality modeling inventory. The base case air quality modeling inventory is used to demonstrate that CAMx can adequately reproduce observed air quality.

In step two, the base case emissions estimates and data are grown and controlled to a future year, the future year emissions estimates and data. In the third step, the future year emissions estimates and data are processed by the emissions model to create the future year air quality modeling inventory. The future year air quality modeling inventory is used in CAMx to establish the baseline air quality field. The baseline air quality field is then used to determine what additional emissions controls, if any, are needed to reach attainment. In the fourth step, additional emissions controls are applied, if necessary, to the future year emissions estimates and data to create the future year air quality modeling controlled inventory. The future year air quality modeling controlled inventory is then used by CAMx to determine if attainment of an air quality metric is reached.

Currently the CARB is developing the base case emissions data and portions of the base case air quality modeling inventory (i.e., biogenics and on-road mobile source emissions estimates). The CARB will develop the state-wide growth and control estimates that are needed to project the base case emissions estimates and data to the future year emissions estimates and data. The CARB will then apply EMS-95 the future year emissions estimates and data to create the future year air quality modeling inventory for the point and area sources. For the on-road mobile sources, the CARB will provide future year gridded, hourly emissions estimates based on the ITN though it is not clear at this point the methods that ARB intends to use to forecast the ITN related data. For the biogenic emissions estimates, it is our understanding that the CARB will use the base year episodic emissions since it is difficult to develop consistent future year land use data sets. As the CARB develops documentation for their methods to develop the future year emissions estimates, we will provide it to the study sponsors.

In coordination with the BAAQMD, the project team will develop the emissions control strategies specific to the SFBA. We will apply those strategies to the future year emissions estimates and data to develop the future year air quality modeling controlled inventory. Though the development of the future year air quality modeling inventory and the future year air quality modeling controlled inventory is an iterative process, we will be limited by project resources in how many air quality modeling inventories can be created.

DELIVERABLES

The deliverables for this task are as follows:

- A report on the development of the base case, future base, and future controlled modeling inventories. The report will be posted to the Contractor-maintained web site. All reports will be provided in both Microsoft Word and Adobe PDF formats.

- The emissions modeling system comprised of all the software components and databases mentioned above. The software components are:
  a) EMS-95 and supporting software systems
  b) CARB models (EMFAC and BEIGIS)
  c) Raw CARB inventories
  d) Ancillary supporting data (surrogates, cross-reference files, etc.)
  e) Processed model-ready inventories used in CAMx
  f) Any supporting components for data conversion
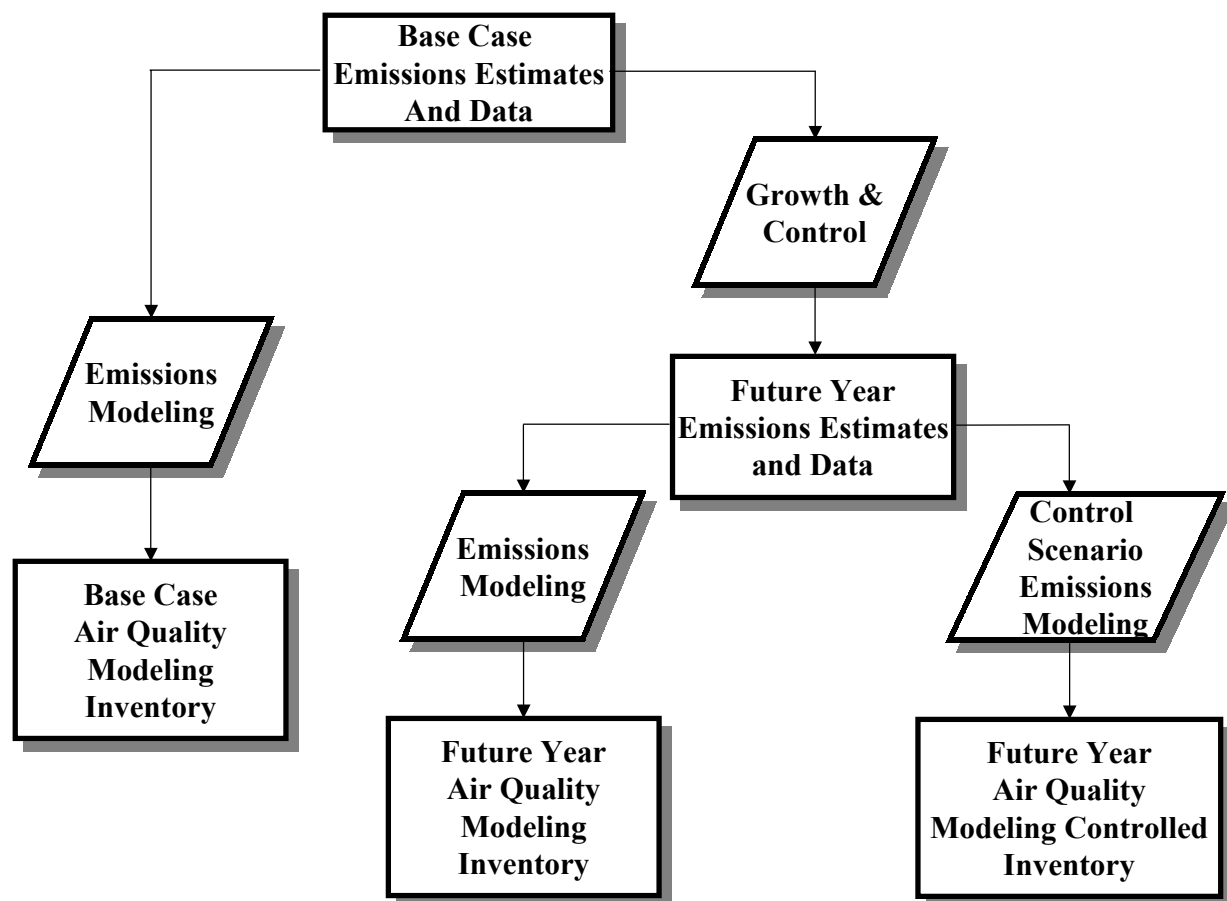
Figure 5-1. Overview of the air quality modeling emissions inventory development.

6. CAMx INPUT DATA PREPARATION

Several data preparation tasks are required to provide CAMx with various inputs that define the meteorology, emissions, initial and boundary conditions, surface characteristics, and photochemical conditions of the atmosphere. The bulk of work associated with meteorology and emissions is described in previous sections. However, some additional processing is needed for these components just before the air quality model is run. This section begins with a discussion on the air quality modeling grid specification; it is critical to define the grid system extent and resolution before the development of most of the CAMx input fields can begin. This section then goes on to describe the procedures to develop and/or format the various input files for CAMx, and finally lists the model options that will be invoked in the base and diagnostic simulations.

CAMx DOMAIN AND GRID SPECIFICATIONS

The spatial domain (or volume) on which Eulerian models operate is defined as a three-dimensional grid, which is used to discretize the environment into averages contained within many small grid cell volumes. The modeling grid should be defined with sufficient size and resolution to capture all of the significant physical processes and transport patterns that affect pollutant concentrations in the focus area. Obviously a balance must be struck between grid size and resolution, both because of resource constraints (budget, schedule, and computing power), and because of limitations inherent in all Eulerian models to characterize physical phenomena at small scales (<1 km horizontally).

Therefore, an important step in the design of an ozone modeling system is specifying the extent of the domain and resolution of the grid. The air quality modeling domain and grid specifications for this study will be based on CARB's current emissions and air quality modeling configuration (mapping projection, domain alignment, etc.). However, through consultation with the District and the MAC, the specific domain extent, nesting configuration, and grid resolution will provide a balance between adequately treating the Bay Area ozone problem and addressing regional transport to and from the SFBA.

The CARB is currently undertaking simulations of the CCOS episodes using MM5 with SAQM, CMAQ, and CAMx, which together with EMS-95 are applied on a very large regional domain on a Lambert Conic Conformal projection with 4-km grid spacing (see Figure 4-1). The RAMS model to be used in this study operates on a Rotated Polar Stereographic projection, and so the RAMSCAMx interface processor will provide the link that performs the necessary manipulations of the RAMS output to properly feed into CAMx on the CCOS Lambert projection. Clearly, the definition of the RAMS polar grid and the CCOS/CAMx Lambert grid will need to be closely coordinated to minimize distortions between the two projections.

Figure 6-1 shows the arrangement of the various CAMx nested grids. The outermost CAMx grid will provide regional coverage and extend over most of the CCOS grid shown in Figure 4-1. In this way, the effects of regional transport into and from the SFBA can be included
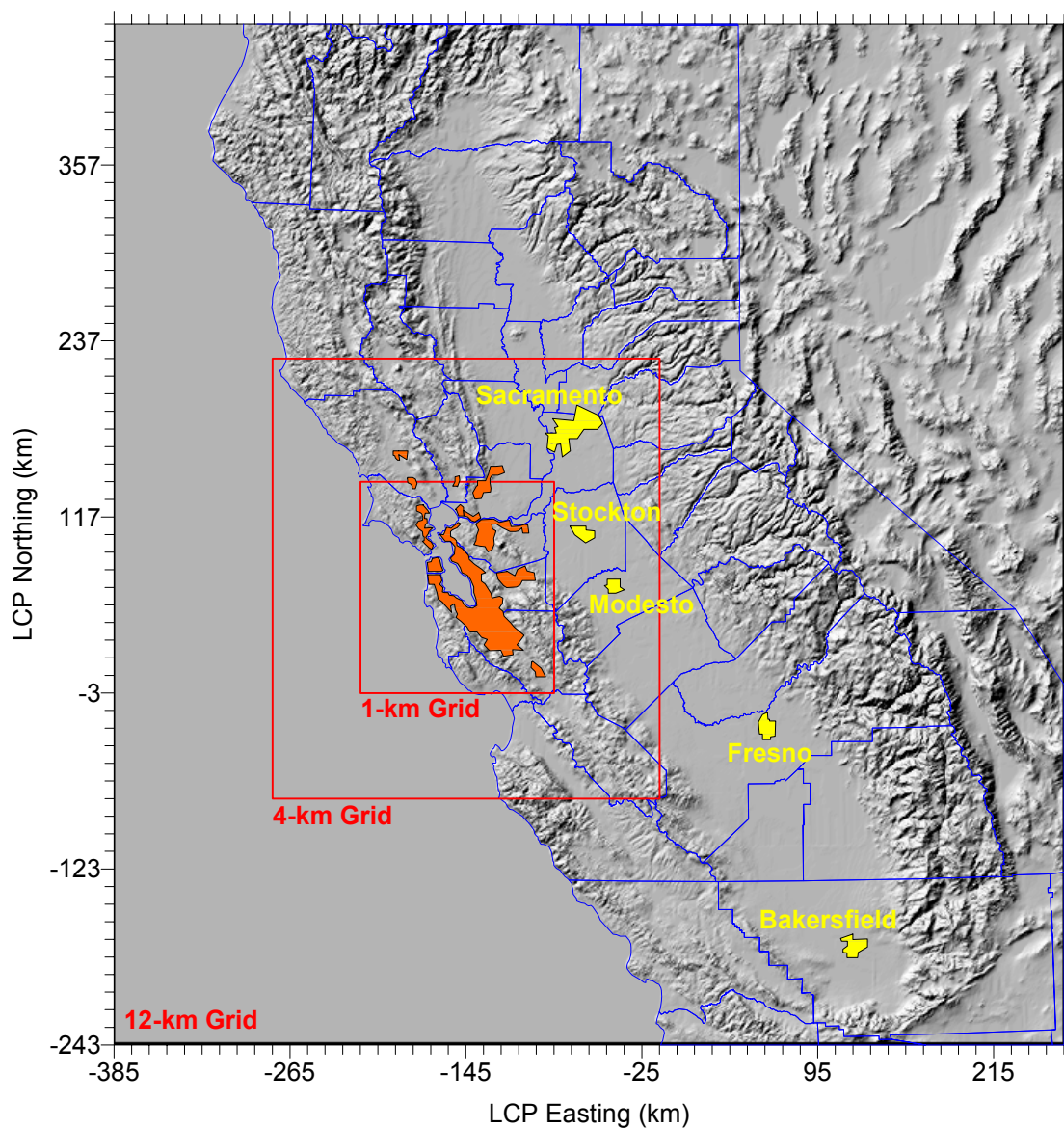
Figure 6-1. Example of the 12/4/1-km CAMx nested grid system covering central California. This domain aligns onto the CCOS 4-km modeling grid.

explicitly in the simulations. Since the emissions inventory will be provided for the CCOS domain, the outermost CAMx grid is limited to the CCOS extent. However, there are two important differences:

(1) While the CCOS grid resolution is 4 km over the entire area, the CAMx grid cell resolution for the outer domain will be set to match the RAMS regional resolution of 12 km (see Figure 4-2, domain 2). This resolution is sufficient to capture the influences of regional air quality on SFBA ozone from the various rural areas outside of central California on the fringe of the CCOS domain. It is also sufficient to resolve the dispersion of SFBA pollutant plumes to the southern San Joaquin Valley (Fresno and Bakersfield) and northern Sacramento Valley as the plumes will be rather diffuse and resolvable at 12 km at such distances downstream. A simple grid cell aggregation step will be necessary to process 4-km emission inputs to the 12-km grid inputs.

(2) The CAMx southern and eastern boundaries are moved inward (north and west, respectively) relative to the CCOS domain. The eastern boundary of the CCOS domain extends well into Nevada, an area that is associated with very little emissions and should not affect the SFBA whatsoever; therefore the eastern boundary of the CAMx grid is placed just east of the Sierra Nevada mountain range. The southern boundary of the CCOS domain extends into the San Fernando Valley of Los Angeles. For CAMx, the southern boundary is moved to a position just south of Tejon Pass and the crest of mountains dividing the South Coast Air Basin and the San Joaquin Valley. Any influx of pollutants from the South Coast Air Basin will be treated via boundary conditions and their potential contribution tested using inert tracer runs and/or the CAMx DDM tool.

The primary inner nest shown in Figure 6-1 will cover the urbanized areas of central California at a grid resolution of 4 km. The purpose of this grid is to provide adequate resolution over the SFBA region and outward to the Districts immediately downwind. The 4-km CAMx nest will exactly overlay a subset of the CCOS grid, so gridded emissions data will simply be "windowed out" for this intermediate grid. The 4-km CAMx nest will be used for most of the CAMx developmental/diagnostic simulations.

Finally, a high-resolution nest will cover the urbanized portion of the immediate SFBA for some sensitivity tests. The resolution is expected to be on the order of 1 km, and will be consistent with the extent of the RAMS high-resolution grid 4 (Figure 4-2). Since topography is a major factor in ozone formation in the SFBA, a 4-km grid will likely not resolve certain wind flow features that may prove critical to the accurate placement and formation of ozone. If tests with the high resolution nest definitively show improved results, then the modeling effort will use this grid for the regulatory simulations. Initial tests will be undertaken with 4-km emissions but with the higher resolved meteorological inputs fields. Gridded emission estimates (area, on-road, biogenics, etc.) will need to be reprocessed to the higher resolution using new spatial surrogates if we are to investigate the full potential effects of this fine grid. Point sources would not need to be reprocessed as those inputs are not dependent upon model resolution.

In the vertical, CAMx will resolve the atmosphere into about 20 layers up to $\sim$7-8 km above the surface (see Figure 6-2 as an example). In past modeling using UAM-V with a 3 km depth, the District has seen large contributions from the top boundary conditions during extended simulations. Also, their past applications of RAMS have generated deep vertical circulation patterns induced by the heights of the coastal ($\sim$1 km) and Sierra ($\sim$3 km) mountain ranges. Thus, there is an apparent need to specify a deeper layer structure than has been typically used in past applications.

CAMx operates in a terrain-following coordinate system, and can match the layer structure of any meteorological model providing three-dimensional gridded input fields. In this case, the CAMx vertical layer structure will be configured to match a subset of RAMS layers. To maximize resolution near the surface and within the typical depth of the daytime boundary layer around the SFBA, CAMx will match most of the RAMS layers within $\sim$1 km of the surface. Above this height, CAMx layers will span multiple RAMS layers. If profiler and ozone sounding data are available during at least one CCOS period, then they will be analyzed to help define the vertical grid structure. Sensitivity tests to domain depth and layer structure will be undertaken to test the model's response to the vertical configuration.

EMISSIONS PROCESSING

While emission files are generated by EMS-95 in model-ready format, there are some final steps to perform before CAMx can be run. First, EMS-95 provides separate gridded surface emission files for each major source category that is processed (i.e., biogenic, area, on-road mobile, etc.). These must be "merged" into a single all-encompassing gridded emissions input file using readily available and standard software tools. Second, the emission files at 4-km resolution must be processed to the various CAMx nested grids. For the outer 12-km grid, the emissions in each 4 km cell will be aggregated to the nine overlying 12-km cells and a new file written. For the smaller CAMx 4-km grid, the emissions within a subset of the emissions grid will be simply extracted to a new file. Both aggregation and windowing are accomplished using a single emissions manipulation program developed by ENVIRON. If the 1-km CAMx grid is utilized without any emissions processing at that resolution, CAMx will automatically map the 4-km emissions to the 1-km grid internally (note that this process does not "resolve" the emissions density any better than 4 km).

Third, large NOx elevated point sources must be selected for the CAMx Plume-in-Grid (PiG) treatment. Point sources will be chosen for the PiG treatment based on a minimum threshold daily NOx emissions rate that yields a few hundred PiG sources in the modeling domain. We propose to apply a lower NOx threshold for sources in and around the focus SFBA (e.g., $\sim$5 tons per day), and higher thresholds in the remainder of California (e.g., $\sim$10 tons per day). We will prepare a list of major NOx emitters selected for PiG treatment and review the list with the District and the MAC prior to any simulations with the PiG module invoked.

```
          RAMS Layers                CAMx Layers
   k      height   thickness         k       thickness
   =======================|  |=======================
   40     7598.7    1000.0        --21---     1951.4
   39     6598.7     951.4
   38     5647.3     792.8        --20---     1453.4
   37     4854.5     660.6
   36     4193.9     550.6        --19---     1009.3
   35     3643.3     458.7
   34     3184.6     382.4        --18---      701.0
   33     2802.2     318.6
   32     2483.6     265.5        --17---      671.1
   31     2218.1     221.2
   30     1996.9     184.4
   29     1812.5     153.7        --16---      388.3
   28     1658.8     128.0
   27     1530.8     106.6
   26     1424.2      96.1        --15---      265.8
   25     1328.1      86.5
   24     1241.6      83.2
   23     1158.4      80.0        --14---      230.9
   22     1078.4      76.9
   21     1001.5      74.0
   20      927.5      71.2        --13---      139.5
   19      856.3      68.3
   18      788.0      65.8        --12---      129.0
   17      722.2      63.2
   16      659.0      60.8        --11---      119.3
   15      598.2      58.5
   14      539.7      56.2        --10---      110.3
   13      483.5      54.1
   12      429.4      50.5        ---9---       97.7
   11      378.9      47.2
   10      331.7      44.1        ---8---       85.4
    9      287.6      41.3
    8      246.3      38.5        ---7---       74.5
    7      207.8      36.0
    6      171.8      33.7        ---6---       33.7
    5      138.1      31.4        ---5---       31.4
    4      106.7      29.4        ---4---       29.4
    3       77.3      27.5        ---3---       27.5
    2       49.8      25.7        ---2---       25.7
    1       24.1      24.1        ---1---       24.1
    0        0.0   ========|  |========Surface======
```

Figure 6-2. RAMS and suggested CAMx vertical grid structure based on 40 sigma-z levels. Heights (m) are above ground level (AGL).

METEOROLOGICAL PROCESSING

Raw output from the RAMS meteorological model needs to be converted to formats and variables used by CAMx specifically. ENVIRON has authored widely used RAMS and MM5 translation software to complete this task. The software includes the ability to interpolate data from the native map projections used by the meteorological models to any projection to be

specified for air quality model (CAMx may be applied on Lambert Conformal, Polar Stereographic, or UTM projections, or in geodetic latitude/longitude). The meteorological translation software will provide an important component of the District's unified RAMS/CAMx modeling system.

CAMx requires meteorological input data for the parameters described in Table 6-1. All of these input data will be derived from the RAMS results. RAMS output fields will be translated to CAMx-ready inputs using ENVIRON's RAMSCAMx translation software. This program performs several functions:

1. Extracts data from the RAMS grids to the corresponding CAMx grids; in this study, the extraction will include a mass-weighted interpolation from the RAMS polar stereographic grid to the CAMx Lambertian grid, with appropriate rotation of vector (wind) variables.
2. Performs mass-weighted vertical aggregation of data for CAMx layers that span multiple RAMS layers.
3. Diagnoses key variables that are not directly output by RAMS (e.g., vertical diffusion coefficients and cloud information).

Table 6-1. CAMx meteorological input data requirements.

CAMx Input Parameter	Description
Layer interface height (m)	3-D gridded time-varying layer heights for the start and end of each hour
Winds (m/s)	3-D gridded wind vectors (u,v) for the start and end of each hour
Temperature (K)	3-D gridded temperature and 2-D gridded surface temperature for the start and end of each hour
Pressure (mb)	3-D gridded pressure for the start and end of each hour
Vertical Diffusivity (m^2/s)	3-D gridded vertical exchange coefficients for each hour
Water Vapor (ppm)	3-D gridded water vapor mixing ratio for each hour
Cloud Cover	3-D gridded cloud cover and liquid water content for each hour
Rainfall Rate (in/hr)	2-D gridded rainfall rate for each hour

The RAMSCAMx program has been written to carefully preserve the consistency of the predicted wind, temperature and pressure fields output by RAMS. This is the key to preparing mass-consistent inputs for CAMx, and therefore for obtaining high quality performance from CAMx.

Care must be exercised in dealing with the different map projections. The RAMS projection has been defined to match the CCOS Lambert projection as closely as possible in the focus area. It is emphasized that simply assigning a cell-by-cell correspondence of meteorological variables to a slightly different air quality grid and/or projection is not an appropriate approach. Ideally, we would want full consistency between the emissions, meteorological and air quality models. However, like CARB, we also believe it is important to be flexible in selecting various model components, and to consider alternative models (e.g., RAMS vs. MM5) that may represent the ambient conditions in the best possible manner.

The data prepared by RAMSCAMx will be directly input to CAMx. Vertical diffusivities (Kv) are an important input to the CAMx simulation since they determine the rate and depth of mixing in the planetary boundary layer (PBL) and above. In general, our experience has been that diffusivities from meteorological models require careful examination before they are used in air quality modeling. This may be because the air quality model results are much more sensitive to diffusivities than the meteorological model results. We will evaluate the CAMx diffusion inputs by comparing the Kv values taken directly from RAMS with several diagnostic calculation approaches, and by analyzing available sounding data from profilers and rawinsondes. Sensitivity simulations will be undertaken with the various Kv fields. Based on prior experience, we will likely apply minimum diffusivity values between layers 1 and 2 to ensure that nocturnal stability near the surface is not over-stated. The minimum value used will depend upon landuse (e.g., urban, forest, agricultural, water, etc.) to represent different impacts of mechanical mixing and surface heat input (e.g., urban heat island effect).

DEVELOPMENT OF ANCILLARY INPUTS

The preparation of ancillary inputs files include initial/boundary conditions, land use distribution for all grids, chemistry parameters, albedo/haze/ozone fields, and photolysis rates.

Initial and Boundary Conditions

The initial conditions (ICs) are the pollutant concentrations specified throughout the modeling domain at the start of the simulation. Boundary conditions (BCs) are the pollutant concentrations specified at the perimeter of the modeling domain. One of the reasons for performing regional scale modeling rather than urban scale modeling is to minimize the importance of ICs and BCs. Using a large regional domain moves the boundaries far away (in distance and transport time) from the study area. Including several "spin-up" days prior to the episode period allows time for the influence of initial conditions to be removed.

As a starting point, the initial and boundary concentrations currently specified by CARB in their CCOS modeling effort will be utilized for CAMx; the CCOS approach will be reviewed and discussed with CARB modeling staff. We expect that certain species' boundary conditions (ozone, NOx) within the mixed layer will be based on surface measurements where available, especially along the southern boundary that divides the southern San Joaquin Valley and the Los Angeles basin. Specification of BC's will also likely include the use of speciated aircraft measurements to specify time/space varying boundary conditions aloft and in remote areas without surface instrumentation. Even if instrumented aircraft did not fly for one or more modeling episodes, we will analyze such data to provide "representative" concentrations in remote areas. In any event, sensitivity runs to investigate the contribution of the boundaries to ozone in the focus area will be undertaken.

Relatively clean and uniform initial conditions will be specified, and are expected to be removed via the addition of a three-day model spinup period before each modeling episode. The effect of the uniform IC assumption, and its contribution to ozone during the focus days, will be analyzed via sensitivity simulations. If the IC contribution is not determined to be

"minimal", then either (1) the spinup period will be lengthened, or (2) spatially varying IC's will be specified based on any available ambient speciated measurement data in the domain. The contribution of initial and boundary conditions to ozone formation in the SFBA may also be investigated through the use of CAMx "Probing Tools" such as Ozone Source Apportionment, the Decoupled Direct Method, or Process Analysis (ENVIRON, 2002), depending upon available resources and schedule.

Surface Characteristics (Landuse)

CAMx requires gridded landuse data to characterize surface boundary conditions, such as roughness, deposition parameters, albedo, vegetative distribution, and water/land boundaries. The land cover categories utilized by CAMx are based on the 11 category system established in RADM, which are parallel with SAQM and UAM-V.

Land use inputs will be developed from three possible sources: (1) emission surrogates used in emissions processing; (2) high-resolution ($\sim$200 m pixel) landuse/landcover data that is freely availble from the USGS in 1:250,000 scale quad maps; and/or (3) any additional high-quality land use data available from local Bay Area planning or other governmental agencies.

ENVIRON and Alpine Geophysics have both developed and possess software to convert raw land cover data to model-ready input variables and formats. These include both Fortran-based systems, and GIS/Arc-Info capabilities. The latter would be crucial if local SFBA data were to be made available to the project.

Chemistry Data

Three input files define the chemistry used in CAMx.

Chemistry Parameters: The chemistry parameters file selects which chemical mechanism to use and specifies the rate constants for the thermochemical reactions. CAMx will be run with the most up-to-date version of the Carbon Bond 4 mechanism (CB4), which is referred to as "mechanism 3" in CAMx. Mechanism 3 is the CB4 mechanism with updated (circa 1995) radical termination reactions and isoprene chemistry as used for the OTAG modeling of the eastern U.S. Alternatively, CAMx may also be run with the SAPRC99 mechanism. While SAPRC99 is newer, provides up-to-date reaction rates, and the hydrocarbon lumping scheme resolves VOC more precisely, to our knowledge it has not definitively performed any better than CB4 in terms of ozone air quality model predictions. SAPRC99 also contains many more reactions and species than CB4, and this leads to model run times are nearly twice that of CB4. Our approach is to use CB4 as the default mechanism in developing the base cases for each episode, and to compare results with SAPRC99 in sensitivity tests. If a clear case can be made for improved model performance in simulating both ozone and precursors over all episodes, then we would adopt SAPRC99 for future year analyses.

Photolysis Rates: The photolysis rates file determines the rates for chemical reactions in the mechanism that are driven by sunlight. The photolysis rates file will be prepared using version 4 of the TUV radiative transfer model developed at NCAR. The rates file is

essentially a very large multi-dimensional lookup table that defines the variation of photolysis reactions over zenith angle, altitude, surface UV albedo, haze turbidity, and total vertically integrated ozone column density.

Albedo/Haze/Ozone File: The albedo/haze/ozone file specifies how these parameters vary in time and space for the CAMx simulation. The photolysis rates and albedo/haze/ozone files must be coordinated to function together correctly. The surface albedo will be calculated based on the gridded landuse data. The stratospheric ozone column data will be based on available satellite data from http://www.cpc.ncep.noaa.gov. Since there may not be a source of regionally specific haze data for the study area, constant haze turbidity representative of rural areas will be assumed over the entire grid. The CCOS database will be reviewed to identify any source of data that may allow for a more robust manner in setting haze turbidity values. Tests with CAMx over a wide but representative range of turbidity values have shown that model results are not particularly sensitive to how this parameter is set.

CAMx MODEL OPTIONS

CAMx has several user-selectable options that are specified for each simulation through the CAMx control file. Most of these options follow naturally from other choices about model inputs. There are three optional inputs that must be decided for this project: the advection scheme, the plume-in-grid scheme and the chemistry solver. The recommended choices for these options are discussed below. See the CAMx User's Guide (ENVIRON, 2002) for more details on these options. The selection for each option will be decided at the stage of the base case model performance evaluation and then held fixed for the remainder of the project.

Advection scheme: CAMx v3.10 has three optional methods for calculating horizontal advection (the movement of pollutants due to horizontal winds) called Smolarkiewicz, Bott and Piecewise Parabolic Method (PPM). The Smolarkiewicz scheme has been used for many years, and was used in many previous studies in California with SAQM and UAM. The Smolarkiewicz scheme has been criticized for causing too much artificial diffusion of pollutants, tending to "smear out" features and artificially overstate transport. The Bott and PPM schemes are newer and have less artificial diffusion than the Smolarkiewicz scheme. Our experience with these schemes suggests that PPM is a better overall approach given that Bott tends to generate some small but definite numerical artifacts. Hence, we are leaning toward using the PPM scheme for this study. On the other hand, the CAMx Decoupled Direct Method (DDM) probing tool is coded to only utilize the Bott approach. Therefore, if DDM is to be extensively used in this project, we may need to establish Bott as the default standard in these applications. On balance, we need to use an overall technical approach that best serves the needs of the study to accurately and efficiently develop a technical defensible ozone attainment plan within the regulatory time table for the SIP development. Sensitivity to both options may be examined in diagnostic tests for the base case.

Plume-in-Grid: CAMx includes an optional sub-grid scale plume model that can be used to represent the dispersion and chemistry of major NOx point source plumes close to the source. We will use the Plume-in-Grid (PiG) sub-model for major NOx sources (i.e., point sources with episode average NOx emissions greater than 5-10 tons per day). Sensitivity to no PIG

treatment may be examined in diagnostic tests for the base case. A new version of PiG may be available in time for use in this project. We will investigate the effects of the new approach in tests.

Chemistry Solver: Starting with version 3 of CAMx there are two options for the numerical solution scheme for the gas phase chemistry. The first option is the CMC fast solver that has been used in every prior version of CAMx. The second option is an IEH solver. The CMC solver is fast and more accurate than most chemistry solvers used in current ozone models. The IEH solver is even more accurate than the CMC solver but significantly slower. The CMC solver will be used for this study. Sensitivity to using the IEH solver may be examined in diagnostic tests for the base case.

7. BASE YEAR MODEL PERFORMANCE EVALUATION

For the base year modeling CAMx will be run for the three historical episodes and the performance of the model will be evaluated against available air quality data. The purpose of the evaluation is to build confidence in the model's reliability as an ozone prediction tool. The proposed evaluation plan will follow the procedures recommended in the EPA and CARB guidance documents for 1-hour ozone (EPA, 1991; CARB, 1992), and new draft guidance for 8-hour ozone (EPA, 1999).

APPROACH TO MODEL PERFORMANCE EVALUATION

It is first important to establish a framework for assessing whether the photochemical modeling system performs with sufficient reliability to justify its use in projecting future year ozone levels and developing ozone control strategies. The framework for assessing the model's reliability consists of the following principals, which are based on EPA's draft 8-hour modeling guidance:

- **The Model Should be Viewed as a System.** When we refer to evaluating a "model" we include not only the CAMx photochemical model, but its various companion preprocessor models, the supporting aerometric and emissions database, and all other related analytical and numerical procedures used to produce modeling results.

- **Model Acceptance is a Continuing Process of Non-Rejection.** Over-reliance on explicit or implied model "acceptance" criteria should be avoided, including EPA's performance goals (EPA, 1991). Models should be accepted gradually as a consequence of successive non-rejections, and confidence builds as the model undergoes a number of different applications (hopefully involving stressful performance testing) without encountering major or fatal flaws that cause the model to be rejected.

- **Criteria for Judging Model Performance Must Remain Flexible.** This approach recognizes the several new elements introduced to SFBA regional application including the use of the latest local-regional emissions data sets and models, the availability of CCOS field study data for two of the episodes, and the lack of special study data for the third.

- **Previous Experience is Used as a Guide for Judging Model Acceptability.** Interpretation of the CAMx modeling results for the episode, against the backdrop of previous modeling experience, will aid in identifying potential performance problems and suggest whether the model should be tested further or rejected.

The specific procedures outlined in the draft 8-hour modeling guidance provide six means by which to establish acceptable model performance:

1. Inspection of computer generated graphics.
2. Calculation of ozone statistical metrics.
3. Comparison of predicted and observed precursor emissions or species concentrations.

4. Comparison of observed and predicted ratios of indicator species.
5. Comparison of predicted source category contribution factors with estimates obtained using observational models.
6. Retrospective analyses in which air quality differences predicted by the model are compared with observed trends.

Sufficient fulfillment of these six points requires the availability of comprehensive measurement data on ozone and precursors from an extensive monitoring network. This may not be feasible in all cases, particularly in regards to precursor measurements. It is also quite possible that the list given above will change with the release of final guidance by EPA. To the extent possible, each of the performance procedures described by EPA's 8-hour guidance will be addressed, and at a minimum, an explanation of why certain components cannot be fulfilled will be provided (e.g., insufficient observational data).

Incorporating the principals listed above into an operational philosophy for judging model performance, we suggest the following approach for assessing the reliability of the CAMx for control strategy development. The evaluation of performance for each of the episodes will be carried out in two sequential phases: a screening analysis and an operational/ mechanistic evaluation. The screening analysis begins with statistical and graphical comparisons of modeled and observed ground-level ozone concentrations, and progresses to potentially more illuminating analyses for precursors, product species, pollutant ratios and groupings if possible given available measurement data. If the screening phase suggests that no obvious flaws or compensating errors exist in the simulation(s), the formal operational phase follows. This activity consists of more in-depth diagnostic analyses and tests, including: (1) evaluation of ozone and precursors against available aircraft data; (2) sensitivity/diagnostic simulations, possibly to include the use of the CAMx probing tools; and (3) corroborative analyses. It is important to note that the individual tests and analyses undertaken during the operational phase must comprise a flexible evaluation approach, given constraints on data, project schedule, and resources. This two-phase approach is described in more detail below.

GRAPHICAL AND STATISTICAL ANALYSES

Initial screening of the CAMx base case ozone predictions will be performed for the modeling episodes in an attempt to identify obviously flawed model simulations and to implement improvements to the model input files in a logical, defensible manner. The screening evaluation will employ ozone performance statistics and plots. Graphical displays will be generated using a combination of several common software packages that the ENVIRON team possess in-house, most of which are readily accessible to the project sponsors. These software packages include the Microsoft Excel, Surfer, and publicly-available PAVE. Examples of the types of graphical displays to be considered for each base case include:

- Ozone time series plots;
- Ground-level ozone isopleths;
- Ozone concentration scatterplots;
- Bias and error statistics stratified by sub-region and by time (day of episode).

The model should produce peak unpaired ozone estimation accuracy, overall bias, and gross error statistics within the approximate ranges of ±15-20%, ±5-15%, and 30-35%, respectively, as recommended by EPA (1991). The statistical calculations will be made for various sub-domains; for example the northern Bay Area, Southern Bay Area, Eastern Bay Area, Monterey Bay Area, Sacramento Valley, and San Joaquin Valley (final selection of sub-domains will be made in consultation with the MAC when initial CAMx simulations are completed). The focus of the statistical analysis will necessarily be on assuring and understanding good performance in the SFBA. However, performance in all sub-regions will need to be satisfactory since particularly poor performance in a given sub-domain outside the SFBA might indicate a systematic or compensating error throughout the model.

If the model's performance is better than all of the ranges given above, the base case would not be rejected unless evidence from any supplemental diagnostic or sensitivity simulations suggest unusual or aberrant behavior. If the base case fails any one of the above general ranges, or is especially poor in a given sub-domain outside of the SFBA, it would become necessary to explain why the performance is poorer than commonly achieved in similar applications and whether the problems will compromise the projection of future ozone levels or the evaluation of emission control strategies. Otherwise, the particular base case in question should be declared inadequate. This outcome would result in one of several courses of action: (a) diagnose the causes of poor performance and rectify such problems, or (b) eliminate the poor-performing episode from use in strategy development and/or identify an alternative episode for substitution in the study.

The graphical and statistical evaluation utilized in the screening evaluation for ozone by sub-domain will also be generated for NO_x and VOC to the extent possible (usually speciated VOC samples are taken over much longer periods at a few sites, so statistics based on the small population set are not particularly useful). Other metrics such as precursor and product ratios will be calculated and compared to measurement data. Should any obvious flaws be detected, model diagnosis and performance improvement efforts may be needed to fully identify and correct (if possible) the noted problems.

Experience in photochemical modeling is the best basis upon which to identify obviously flawed simulation results. Efforts to improve photochemical model performance, where necessary and warranted (i.e., to reduce the discrepancies between model estimates and observations), should be based on sound scientific principles. A "curve-fitting" or "tuning" activity is to be avoided. The following principals should govern the model performance improvement process (to the fullest extent possible given the project schedule):

- Any significant changes to the model or its inputs must be documented;
- Any significant changes to the model or its inputs must be supported by scientific evidence, analysis of new data, or by re-analysis of the existing data where errors or misjudgments may have occurred; and
- All significant changes to the model or its inputs should be reviewed by the MAC.

OPERATIONAL/MECHANISTIC EVALUATION

Performance Evaluation Against Aircraft Data

The air quality and meteorological data compiled from standard and special study networks such as CCOS can be used in "stand-alone" analyses such as developing conceptual models for ozone formation. They are also valuable for improving photochemical modeling studies to provide more reliable air quality plans. Aircraft data are of particular value because they provide data away from the fixed surface monitoring network (i.e., aloft and between surface sites) and so can be useful in evaluating boundary conditions, individual pollutant plumes, and transport of regional sources to urban areas of interest.

ENVIRON has used aircraft data in several model performance evaluations in Texas. We have developed comparison methodologies and software that take into account the different temporal and spatial scales represented by aircraft data and photochemical model results.

As stated in Section 2, four instrumented aircraft were used to measure the vertical and horizontal gradients of temperature, humidity, and pollutant concentrations in the study region during CCOS IOPs. These aircraft included a Cessna 172RG and Cessna 182 operated by University of California at Davis (UCD), and a Cessna 182 and Piper Aztec operated by Sonoma Technology, Inc. (STI). One additional aircraft (Twin Otter), flown by the Tennessee Valley Authority (TVA), made measurements in power plant plumes. The TVA data were collected to evaluate the plume-in-grid parameterizations used in air quality models.

Aircraft data will be procured from the CARB for the specific episodes addressed in this modeling study. These will be used to directly evaluate model performance for ozone and precursors along the flight trajectory paths. Such analyses provide insight into placement of urban and industrial plumes, height of the well-mixed boundary layer, and chemical production aloft (if product species were measured).

Another use for aircraft data is to provide information on setting lateral and top boundary conditions for remote areas such as over the ocean where no routinely available measurements are made. If aircraft flights were not conducted in remote areas during the modeling episodes, we will evaluate whether the data from other flights could be used to set boundary conditions, e.g., similar meteorological/transport conditions, etc.

Diagnostic and Sensitivity Simulations

Objectives

A limited number of diagnostic simulations will be performed to help understand and possibly improve base case model performance. In addition, sensitivity tests will be performed to diagnose model sensitivity to changes in key inputs. These tests are an important component of the base case model evaluation process. In general, diagnostic and sensitivity analyses serve to:

- Reveal model responses that are inconsistent with expectations or other model responses.

- Identify what parameters (or inputs) dominate (or do not dominate) model results.
- Examine the relationship between uncertainties in model inputs and model outputs (error propagation through the model).
- Identify alternate base cases that offer similar model performance and therefore identify potential compensating errors.
- Provide guidance for model refinement and data collection programs.

The exact sensitivity simulations that will be needed can only be assessed after the initial model evaluations are performed.

Tests That Are Not Recommended

With the advent of more sophisticated nested regional ozone models (such as CAMx) a number of sensitivity runs that were historically carried out with the UAM and other models are no longer needed or appropriate. These tests include zero-emission, zero initial condition, zero boundary condition runs and modified wind field tests such as halving the wind speeds. Physically unrealistic tests such as these can produce misleading results that are difficult to interpret. For the zero emission and zero IC/BC sensitivity tests, more can be learned from looking at sensitivity to alternate (but physically possible) inputs. Ad-hoc modifications to wind fields external to meteorological models like RAMS and MM5 are not recommended because they destroy consistency among the meteorological inputs (e.g., winds that are physically unrelated to pressures and temperatures). Other types of meteorological experiments are potentially more useful, such as alternate vertical eddy diffusivities or alternate vertical grid structures.

Recommended Tests

Sensitivity experiments will be considered as part of the performance evaluation analysis as appropriate. The potential need for and nature of these simulations will be discussed with the District, CARB, and MAC representatives. Up to 10 sensitivity/diagnostic CAMx simulations are planned.

Potential diagnostic runs include:
- Initial and boundary conditions; examine the extent to which they effect or contribute to the ozone simulated in the focus area and downwind.
- Biogenic emissions; evaluate potential effects of uncertainties in biogenic emissions levels.
- PiG treatment; check how the implementation of this submodel affects ozone patterns.
- Advection scheme; quantify the impact of various options on performance.
- Meteorology; design specific diagnostic tests identified during the preparation of the meteorological inputs such as: alternate vertical diffusion coefficients; impacts of clouds on photolysis rates; effects of grid resolution (horizontal and vertical); and effects of alternative meteorological realization from varying nudging strengths and/or various model options.
- Chemistry; check the influence of the mechanisms (CB4 vs. SAPRC99) and the specific numerical solvers (CMC vs. IEH).

Potential sensitivity runs include:

- Sensitivity to modified initial and boundary concentrations, as they may have an effect on assumptions in extrapolating to future year conditions.
- Sensitivity to reductions/increases in total anthropogenic VOC and/or NOx emissions.
- Sensitivity to reductions/increases in anthropogenic VOC and/or NOx emissions from specific source categories.
- Sensitivity to reductions/increases in anthropogenic VOC and/or NOx emissions from specific source regions (e.g., distant or local).

Use of CAMx Probing Tools for Diagnostic Evaluation

CAMx provides several "extensions" to the basic chemical/dispersion model, referred to as "probing tools", that provide information concerning source apportionment and the relative importance of various physical and chemical processes. These tools include the Ozone Source Apportionment Technology (OSAT and derivatives), Process Analysis (PA), and the Decoupled Direct Method of tracking sensitivity coefficients (DDM, similar to a source apportionment of emissions). All of these are described in the CAMx User's Guide for version 3.10.

For diagnostic purposes, the most useful tools are DDM and PA, which provide a wealth of information concerning the rates of change in ozone relative to transport and chemical processes. Reviewing this information can lead to insights into specific model performance issues and NOx/VOC-limited chemical kinetics in user-specified portions of the domain. OSAT is typically used for assessing source apportionment for purposes of designing control strategy scenarios, but can be used like DDM in a diagnostic manner to assess the relative importance of various sources. In this way, the user may be able to discover if a particular source area/category is having a stronger or weaker influence on ozone in key receptor areas than conceptually expected, and then undertake an investigation to determine if that signal is appropriate or not.

Depending on project resources and schedule, use of these probing tools will be considered in the diagnostic evaluation of the CAMx Base Case simulations.

Corroborative Analyses

Recently, emphasis has been continuously placed upon utilizing independent air quality and/or emissions analyses to corroborate the findings of air quality model simulations. Draft 8-hour ozone and PM modeling guidance from EPA now stress the need to evaluate air quality model performance and estimated source attribution against such independent examinations in "weight of evidence" analyses, which are used to support or refute the signals obtained from the simulations. The most useful and varied corroborative examinations are available to an air quality modeling study when it is coordinated with a multi-faceted field study program.

Fortunately, the data analyses planned for the CCOS period could provide ample benefits to this modeling study. The study team and District staff will monitor the progress of the CCOS

data analysis projects through spring of 2003 and note any particular analyses that could help explain model performance or guide control strategy development. Depending on available project resources and schedule at the time, the most relevant of these will be considered for corroborative model evaluation. Particularly relevant and useful analyses would include NOx/VOC limited chemistry, Chemical Mass Balance (CMB) modeling for hydrocarbons, any retrospective trends analyses for NOx, VOC and ozone, and Blanchard's MAPPER results for sites in and around the SFBA. We will also compare CAMx base case results with any of CARB's modeling results when they are made available.

DELIVERABLES

The deliverables for this task are as follows:

- A report on the development, configuration, and evaluations of the air quality simulations of the three modeling episodes. The report will be posted to the Contractor-maintained web site. All reports will be provided in both Microsoft Word and Adobe PDF formats.

- The base year air quality modeling system comprised of all the databases and software components developed to support the model performance evaluation.

8. FUTURE/ATTAINMENT YEAR OZONE MODELING

The ozone model developed under this protocol will be used for evaluating the effectiveness of future emissions control scenarios for the SFBA. It will be necessary to show that the 2006 SFBA emissions inventory will be sufficient to reduce the current ozone design value to below the 124 ppb NAAQS. The methods used for this activity will be consistent with current 1-hour (EPA, 1991; 1996) and evolving 8-hour guidance (EPA, 1999).   The evaluation of control strategies will be carried out under the direction of the District and the MAC.

The CAMx future-year baseline simulations will be helpful in assessing the extent to which further emissions reductions are needed in the region to provide for attainment of the 1-hour ozone NAAQS in the SFBA. The strictest test for attainment is the "deterministic" approach, in which all grid cells within the regulated area show 1-hour ozone concentrations below 124 ppb for a given package of control measures. However, numerical models are an imperfect representation of reality, with over and under prediction biases and limited precision. Furthermore, the chosen episode(s) may be characterized by ozone levels that are higher or lower than the current design value. Recognizing this, the EPA promotes the use of "weight of evidence" analyses to support the conclusion that a set of selected control measures do provide, by all available evidence, attainment of the ozone NAAQS.

In particular, EPA has developed a methodology called "design value scaling" for using model estimates in conjunction with recent air quality data to estimate whether the ozone standard will be attained by a specific future year. Briefly, the current ozone design value is scaled by the model predicted change in ozone levels between the base and future years. If this scaled design value continues to exceed the standard, then additional emission reductions are needed. The control strategy developed by this design value scaling can form the basis for an ozone attainment demonstration. Other types of analyses (e.g., emissions and air quality trends, emission shortfall calculations) can also be included in a weight of evidence argument to support a particular control strategy, as described in the previous section. Since this modeling protocol deals with photochemical ozone modeling, this section focuses on the modeling aspects of an attainment demonstration rather than weight of evidence arguments.

FUTURE YEAR BASE CASE OZONE

A future year of 2006 will be used in this study as that is the new attainment date for the SFBA, according to the 2001 EPA partial disapproval of the previous SIP submission. A future year base case will be prepared by adjusting specific model inputs to reflect expected changes between the base (2000) and future (2006) years. The only model inputs that may be changed for the future year are the emissions, initial conditions and boundary conditions.

The future year baseline emissions inventory will be developed by applying source-category specific growth and control factors applied to the 2000 base year inventory. This process is described in Section 5 of this protocol. The initial conditions will be set to relatively clean conditions and allowed to be removed over the course of a three-day model spin up period prior to the days of interest of each episode. Therefore, initial conditions for the future year will remain unchanged from base year values. It is anticipated that the base year boundary

conditions will be set to relatively clean values for remote regions over the Pacific Ocean, near the California-Oregon border, and over Nevada, and that these inputs will have little impact on model results for the SFBA. Therefore, it is expected that these boundary conditions will also be unchanged between the base year and future year values.

However, boundary conditions for the southern boundary, which crosses along the mountains separating the San Joaquin Valley and the South Coast Air Basin, will characterize the urban air flux from southern California. The southern boundary conditions will therefore require future year projection adjustments above the background concentrations assumed for the other boundaries. These adjustment will be estimated from information obtained from the CARB and the South Coast Air Quality Management District.

The results from the future year base case will be analyzed to determine the level and types of additional controls (if necessary) to ensure that the SFBA reaches attainment of the 1-hour ozone NAAQS. Numerous analyses will be considered for this purpose, subject to time and budget constraints, to include:

- NOx/VOC Matrix runs: The future year base case will be rerun with incremental NOx and VOC emission reductions applied uniformly "across-the-board" (by category and over space and time). The emission reductions will range from 0% to 50% in increments of 10%. The resulting peak 1-hour ozone values at 2 to 3 sites will be plotted as in an "Ekma" diagram to help determine the most efficient route (NOx or VOC) to attainment. The sites for plotting will be selected by the District, but should include those historically exceeding the standard (such as Livermore, Gilroy, and Los Gatos).

- Probing Tools: The CAMx OSAT or DDM extensions will be utilized to provide source attribution analyses for the future year base case. These tools provide estimates on the fraction of ozone produced from emissions stratified by source category and geographic area (e.g., on-road mobile in the north bay counties) and indicate whether that ozone is produced from NOx- or VOC-limited chemistry. The OSAT "Ozone Tool" is an Excel-based postprocessor that allows for a complete analysis of the relative contributions from all sources/areas at user-selected sites within the domain (see ENVIRON, 2002).

- Emission Sensitivity Tests: As described in the previous section, numerous sensitivity tests may be carried out with specific emission sectors reduced. These tests would use the results from the matrix and/or probing tool applications listed above to test the effects of targeted emission controls. Results will help to shape the selection of SIP control measures. The effects of these tests on ozone and precursor transport contributions into downwind air quality districts will be analyzed over the entire grid system.

FUTURE YEAR CONTROL/ATTAINMENT OZONE

This District will be responsible to develop the specific emission control measures needed to demonstrate attainment for the 2004 SIP revision. The selection of which sectors to focus on, and what level reduction to reasonably apply, will be guided in part by the modeling analyses described above. Several modeling iterations will likely be necessary with various control

packages and contingencies. The goal of the future year control runs will be to show that simulated ozone is reduced to below 124 ppb over the entire SFBA in 2006. This is often referred to as the "deterministic" test. The effect of the selected control measures on ozone and precursor transport contributions into downwind air quality districts will also be analyzed over the entire grid system.

Additionally, design value scaling may be employed as part of a "weight of evidence" determination that demonstrates the relative reduction of the current ozone design value to below 124 ppb. However, such an approach needs careful consideration, and to the extent possible must address many technical issues early in the process, such as:

- How a peak model prediction away from monitoring sites is treated in the scaling;
- What to do if the design and observed values at a monitoring site are significantly different;
- How poor model performance at some monitoring sites affect the design value scaling.

Whether design value scaling is used or not will depend on these and other issues. One-hour ozone design value scaling has been used in many recent ozone SIPs approved by the U.S. EPA as part of their Weight of Evidence (WOE) attainment demonstration (e.g., Dallas-Fort Worth, St. Louis, Houston/Galveston). This is an evolving area of technical dialog, including guidance at the national level from U.S EPA for attainment demonstrations throughout the U.S. Further, the CARB is certainly facing similar issues in their ozone SIP attainment and maintenance demonstrations for the San Joaquin Valley, South Coast Air Basin, and the Sacramento regional area. As further technical information and guidance becomes available by way of the U.S. EPA, CARB, and the insight of the MAC, we will certainly reflect that combined wisdom in this study.

DELIVERABLES

The deliverables for this task are as follows:

- A report on the development and evaluations of the future year base case, sensitivity, and control strategy air quality simulations of the three modeling episodes. The report will be posted to the Contractor-maintained web site. All reports will be provided in both Microsoft Word and Adobe PDF formats.

- The future year base and control/attainment air quality modeling databases.

Appendix A

EMS-95 Emissions Dataset Formats

Section A-1. Point Source Data File Formats

EMS-95 will be used in the Bay Area SIP revision air quality modeling study. The CARB will deliver EMS-95 derived emissions estimates that are stored in the $EMS_RUN or $EMS directories. The CARB will supply emissions estimates in the formats described below. The following definitions apply to all emissions data file formats that are presented in the protocol (note that units are supplied where necessary in each of the data file formats):

```
>       VARIABLE    -- the name of the EMS-95 variable
>       TYPE        -- the variable declaration
                        I = any valid integer number
                        A = character up to FORMAT in length
                        R = any valid real number
>       FORMAT      -- number of positions the variable has in each record of the data
                        file
>       DESCRIPTION -- brief description of the meaning of each VARIABLE
>       REQUIRED    -- is this a required field?
                        Y = yes, the data is required for proper operation
                        N = no, the data is not required
                        D = it is desirable to have the data, but it is not required
                             for proper operation
```

Also note, that in all cases where stid is required, a value of "6" must be supplied since "6" is the FIPS state identifier for California. The FIPS county identifiers, cyid, for California are as follows:

FIPS Id	Name	FIPS Id	Name	FIPS Id	Name	FIPS Id	Name	FIPS Id	Name	FIPS Id	Name
1	Alameda	21	Glenn	41	Marin	61	Placer	81	San Mateo	101	Sutter
3	Alpine	23	Humboldt	43	Mariposa	63	Plumas	83	Santa Barbara	103	Tehama
5	Amador	25	Imperial	45	Mendocino	65	Riverside	85	Santa Clara	105	Trinity
7	Butte	27	Inyo	47	Merced	67	Sacramento	87	Santa Cruz	107	Tulare
9	Calaveras	29	Kern	49	Modoc	69	San Benito	89	Shasta	109	Tuolumne
11	Colusa	31	Kings	51	Mono	71	San Bernardino	91	Sierra	111	Ventura
13	Contra Costa	33	Lake	53	Monterey	73	San Diego	93	Siskiyou	113	Yolo
15	Del Norte	35	Lassen	55	Napa	75	San Francisco	95	Solano	115	Yuba
17	El Dorado	37	Los Angeles	57	Nevada	77	San Joaquin	97	Sonoma		
19	Fresno	39	Madera	59	Orange	79	San Luis Obispo	99	Stanislaus		

EMS-95 requires five foundation data files in SAS format (from the $EMS directory) for the stationary source emissions data: facility; stack; device; process; and emission; and a day-specific emissions estimates SAS data set (for the $EMS_RUN directories) for each day of the episode. The formats of these SAS data sets are as follows:

$EMS/facility

```
variable type format description                                      required
-------- ---- ------ -------------------------------------------------- --------
stid     I           FIPS state                                         Y
cyid     I           FIPS county code                                   Y
fcid     A    15     facility ID                                        Y
sic      A    4      standard industrial classification                 D
utmx     R           UTM easting (meters)                               D
utmy     R           UTM northing (meters)                              D
utmz     I           UTM zone                                           D
name     A    40     facility name                                      D
emistype A    2      emission type (always PT)                          Y
```

If the UTM coordinates are not supplied, the facility will be placed in the center of the county for purposes of modeling. If the UTM coordinates are supplied and a spatial check reveals that the source is outside of the county, the facility is again placed in the center of the county for purposes of modeling. In either case, EMS-95 flags the facility as incorrectly located.

$EMS/stack

```
variable type format description                                            required
-------- ---- ------ --------------------------------------------------- --------
stid     I           FIPS state code                                        Y
cyid     I           FIPS county code                                       Y
fcid     A    15     facility ID                                            Y
stkid    A    12     stack ID                                               Y
diam     R           inside stack diameter (feet)                           Y
heit     R           stack height above ground surface (feet)               Y
temp     R           stack exit temperature (°F)                            Y
veloc    R           stack exit velocity (feet/second)                      D
flow     R           stack exit flow rate (actual cubic feet/minute)        Y
utmx     R           UTM easting (meters)                                   D
utmy     R           UTM northing (meters)                                  D
elev     R           elevation of stack base from mean sea level (feet)     N
```

If the UTM coordinates are not supplied, the stack will be placed at the coordinates specified for the facility in the facility.pt data file for purposes of modeling. If the UTM coordinates are supplied and a spatial check reveals that the stack is outside of the county, the stack is placed at the coordinates specified for the facility in the facility.pt data file for purposes of modeling. In either case, EMS-95 flags the stack as incorrectly located.

Stack parameters are very critical to the placement of emissions in the vertical air quality modeling domain. It may surprise some that incorrect parameters are sometimes the single largest source of error in the emissions and subsequent air quality modeling, and in some instances, call into question the results of an entire air quality modeling study. EMS-95 will generate a report of suspect stack parameters that will be returned to the ARB for resolution.

$EMS/device

```
variable type format description                                            required
-------- ---- ------ --------------------------------------------------- --------
stid     I           FIPS state code                                        Y
cyid     I           FIPS county code                                       Y
fcid     A    15     facility ID                                            Y
stkid    A    12     stack ID                                               Y
dvid     A    12     device ID                                              Y
sic      A    4      standard industrial classification                    D
dec      R           fractional December throughput                        D
jan      R           fractional January throughput                         D
feb      R           fractional February throughput                        D
mar      R           fractional March throughput                           D
apr      R           fractional April throughput                           D
may      R           fractional May throughput                             D
jun      R           fractional June throughput                            D
jul      R           fractional July throughput                            D
aug      R           fractional August throughput                          D
sep      R           fractional September throughput                       D
oct      R           fractional October throughput                         D
nov      R           fractional November throughput                        D
win      R           winter throughput (Dec - Feb) (%)                     D
spr      R           spring throughput (Mar - May) (%)                     D
sum      R           summer throughput (Jun - Aug) (%)                     D
```

```
fal        R            fall throughput (Sep - Nov) (%)                        D
hours      I            code value for hourly operation                        Y
days       I            code value for daily operation                         Y
weeks      I            weeks of operation per year (weeks/year)               D
dayyear    I            days of operation per year (days/year)                 D
houryear   I            hours of operation per year (hours/year)               D
emistype   A  2         emission type (always PT)                              Y
```

The data in $EMS/device are used to temporalize the emissions estimates (i.e. allocate average
ozone season day emissions estimates into hour-by-hour bins for use in air quality modeling).
The hierarchy for how EMS-95 computes the temporalization factors is beyond the scope of
this document, though, the reader is referred to Wilkinson et al. (1994) for a full description
of how the temporalization factors are estimated. However, if no temporalization data are
supplied, EMS-95 will treat the source as if it emits twenty-four hours per day, 365 days per
year. Of final note, the *hours* and *days* fields are code values that are used to lookup the
actual hours of operation during the day and actual days of operation per week. That is, a
value of 8 for *hours* refers to coded value eight in the appropriate lookup table. The value in
the lookup table indicates that the source operates from 7:00 AM to 3:00 PM local time and
does not operate at any other time during the day.

$EMS/process

```
variable type format description                                           required
-------- ---- ------ -------------------------------------------------- --------
stid     I            FIPS state code                                        Y
cyid     I            FIPS county code                                       Y
fcid     A  15        facility ID                                            Y
stkid    A  12        stack ID                                               Y
dvid     A  12        device ID                                              Y
prid     A  12        process ID                                             Y
scc      A  8         source classification code                            Y
prrt     R            annual process rate (SCC units/year)                   N
prun     A  15        optional process rate units if different from SCC     N
emistype A  2         emission type (always PT)                              Y
```

EMS-95 uses the SCC to select an appropriate TOG speciation profile. If no SCC is supplied,
EMS-95 selects the default TOG speciation profile that in all likelihood does not represent the
actual TOG emissions profile from the source.

$EMS/emission

```
variable type format description                                           required
-------- ---- ------ -------------------------------------------------- --------
stid     I            FIPS state code                                        Y
cyid     I            FIPS county code                                       Y
fcid     A  15        facility ID                                            Y
stkid    A  12        stack ID                                               Y
dvid     A  12        device ID                                              Y
prid     A  12        process ID                                             Y
polid    A  5         pollutant ID                                           Y
                      -- TOG for total organic gases
                      -- NOX for oxides of nitrogen
                      -- CO  for carbon monoxide
acef     R            actual emission factor (tons/SCC units)                N
alef     R            allowable emission factor (tons/SCC units)             N
acee     R            actual emissions (tons/temporal basis)                 Y
alee     R            allowable emissions (tons)                             N
estt     A  2         temporal basis (AA, AD)                                Y
pcec     A  5         primary control equipment                              N
```

```
scec      A   5      secondary control equipment                      N
ceef      R          control equipment efficiency (%)                 N
emistype  A   2      emission type (always PT)                        Y
estt      A       2  estimate type                                    Y
                     -- DS for day-specific emissions estimate
                     -- AD for average season daily emissions estimate
                     -- AA for annual average daily emissions estimate
```

The *polid* field will contain one of the following: CO, NOX, or TOG. The definitions of
VOC, TOG, and ROG are provided in Appendix G. The *estt* field will contain AD for
average ozone season weekday emissions, AA for annual average daily emissions, or DS for
day-specific emissions.

$EMS_RUN/ptemis

```
variable type format description                                          required
-------- ---- ------ --------------------------------------------------- --------
STID      I          FIPS state identifier                                Y
CYID      I          FIPS county identifier                               Y
FCID      A   15     facility ID                                          Y
STKID     A   12     stack ID                                             Y
DVID      A   12     device ID                                            Y
PRID      A   12     process ID                                           Y
POLID     A    5     pollutant ID                                         Y
                     -- TOG for total organic gases
                     -- NOX for oxides of nitrogen
                     -- CO  for carbon monoxide
SIC       A    4     standard industrial classification                   D
SCC       A    8     source classification code                           Y
ICELL     I          east-west cell location in CCOS domain                Y
JCELL     I          north-south cell location in CCOS domain              Y
ACEEBS    R          base emissions estimate (kg)                         Y
ACEEKG    R          actual daily emissions estimate (kg)                 Y
                     -- sum of hremis01-hremis24
ALEEKG    R          allowable daily emissions estimate (kg)              N
HREMIS01  R          hour 0000 - 0100 emissions estimate (kg)             Y
HREMIS02  R          hour 0100 - 0200 emissions estimate (kg)             Y
HREMIS03  R          hour 0200 - 0300 emissions estimate (kg)             Y
HREMIS04  R          hour 0300 - 0400 emissions estimate (kg)             Y
HREMIS05  R          hour 0400 - 0500 emissions estimate (kg)             Y
HREMIS06  R          hour 0500 - 0600 emissions estimate (kg)             Y
HREMIS07  R          hour 0600 - 0700 emissions estimate (kg)             Y
HREMIS08  R          hour 0700 - 0800 emissions estimate (kg)             Y
HREMIS09  R          hour 0800 - 0900 emissions estimate (kg)             Y
HREMIS10  R          hour 0900 - 1000 emissions estimate (kg)             Y
HREMIS11  R          hour 1000 - 1100 emissions estimate (kg)             Y
HREMIS12  R          hour 1100 - 1200 emissions estimate (kg)             Y
HREMIS13  R          hour 1200 - 1300 emissions estimate (kg)             Y
HREMIS14  R          hour 1300 - 1400 emissions estimate (kg)             Y
HREMIS15  R          hour 1400 - 1500 emissions estimate (kg)             Y
HREMIS16  R          hour 1500 - 1600 emissions estimate (kg)             Y
HREMIS17  R          hour 1600 - 1700 emissions estimate (kg)             Y
HREMIS18  R          hour 1700 - 1800 emissions estimate (kg)             Y
HREMIS19  R          hour 1800 - 1900 emissions estimate (kg)             Y
HREMIS20  R          hour 1900 - 2000 emissions estimate (kg)             Y
HREMIS21  R          hour 2000 - 2100 emissions estimate (kg)             Y
HREMIS22  R          hour 2100 - 2200 emissions estimate (kg)             Y
HREMIS23  R          hour 2200 - 2300 emissions estimate (kg)             Y
HREMIS24  R          hour 2300 - 2400 emissions estimate (kg)             Y
MNTH_FAC  R          monthly temporalization factor                       Y
WEEKDAYS  R          weekday temporalization code                         N
DAY_FAC   R          daily temporalization factor                         Y
WEEK_FAC  R          weekly temporalization factor                        Y
PROJ_FAC  R          point source projection factor                       D
PTGF      R          point source growth factor                           D
CNTLCODE  A   15     emissions control code                               D
GRWCODE   A    8     emisisons growth code                                D
```

```
EMISTYPE    A      2   emission type (always PT)                          Y
ESTT        A      2   estimate type                                      Y
                       -- DS for day-specific emissions estimate
                       -- AD for average season daily emissions estimate
                       -- AA for annual average daily emissions estimate
```

Section A-2. Area Source Data File Formats

EMS-95 requires two foundation data files in SAS format (from the $EMS directory) for the area source emissions data: areatprl; and area, and a day-specific emissions estimates SAS data set (from the $EMS_RUN directories) for each day of the episode. The formats of these SAS data sets are as follows:

$EMS/areatprl

```
variable type format description                                         required
-------- ---- ------ ----------------------------------------------------- --------
stid     I           FIPS state code                                       Y
cyid     I           FIPS county code                                      Y
asct     A    15     area source category                                  Y
prrt     R           annual process rate (process units/year)              N
acun     A    15     activity units                                        N
dec      R           fractional December throughput                        D
jan      R           fractional January throughput                         D
feb      R           fractional February throughput                        D
mar      R           fractional March throughput                           D
apr      R           fractional April throughput                           D
may      R           fractional May throughput                             D
jun      R           fractional June throughput                            D
jul      R           fractional July throughput                            D
aug      R           fractional August throughput                          D
sep      R           fractional September throughput                       D
oct      R           fractional October throughput                         D
nov      R           fractional November throughput                        D
win      R           winter throughput (Dec - Feb) (%)                     D
spr      R           spring throughput (Mar - May) (%)                     D
sum      R           summer throughput (Jun - Aug) (%)                     D
fal      R           fall throughput (Sep - Nov) (%)                       D
hours    I           code value for hourly operation                       Y
days     I           code value for daily operation                        Y
weeks    I           weeks of operation per year (weeks/year)              D
dayyear  I           days of operation per year (days/year)                D
houryear I           hours of operation per year (hours/year)              D
emistype A    2      emission type (always PT)                             Y
```

The data in areatprl.ar are used to temporalize the area source emissions estimates (i.e. allocate average ozone season day emissions estimates into hour-by-hour bins for use in air quality modeling). The hierarchy for how EMS-95 computes the temporalization factors is beyond the scope of this document, though, the reader is referred to Wilkinson et al. (1994) for a full description of how the temporalization factors are estimated. However, if no temporalization data are supplied, EMS-95 will treat the area source as if it emits twenty-four hours per day, 365 days per year. Of final note, the *hours* and *days* fields are code values that are used to lookup the actual hours of operation during the day and actual days of operation per week. That is, a value of 8 for *hours* refers to coded value eight in the appropriate lookup table. The value in the lookup table indicates that the source operates from 7:00 AM to 3:00 PM local time and does not operate at any other time during the day.

$EMS/area

```
variable type format description                                         required
-------- ---- ------ ----------------------------------------------------- --------
stid     I           FIPS state code                                       Y
cyid     I           FIPS county code                                      Y
asct     A    15     area source category                                  Y
```

```
polid      A  5      pollutant ID                                    Y
                     -- TOG for total organic gases
                     -- NOX for oxides of nitrogen
                     -- CO  for carbon monoxide
acef       R         actual emission factor (tons/process unit)      N
alef       R         allowable emission factor (tons/proc.unit)      N
acee       R         actual emissions (tons/temporal basis)          Y
alee       R         allowable emissions (tons)                      N
pcec       A  5      primary control equipment                       N
scec       A  5      secondary control equipment                     N
ceef       R         control equipment efficiency (%)                N
emistype   A  2      emission type (always AR)                       Y
estt       A  2      estimate type                                   Y
                     -- DS for day-specific emissions estimate
                     -- AD for average season daily emissions estimate
                     -- AA for annual average daily emissions estimate
```

The *polid* field will contain one of the following values: CO, NOX, or TOG. The definitions of VOC, TOG, and ROG are provided in Appendix G. The *estt* field will contain AD for average ozone season weekday emissions, AA for annual average daily emissions, or DS for day-specific emissions.

Section A-3. Day-specific Biogenic Emissions Estimates

Biogenic emissions estimates will be supplied for each day in each episode. Each day of the episode will have one SAS data set from the $EMS_RUN directory in the following format:

bioemis

```
variable type format description                                            required
-------- ---- ------ --------------------------------------------------- --------
stid     I           FIPS state code                                        D
cyid     I           FIPS county code                                       D
asct     A    15     area source category (always BI)                       Y
icell    I           east-west grid cell number in the domain               Y
jcell    I           north-south grid cell number in the domain             Y
polid    A    5      pollutant ID                                           Y
                     -- ISOP for isoprene
                     -- TERP for monoterpenes
                     -- OVOC for other volatile organic compounds
                     -- MBO  for 2-methyl-3-butenol
adjest01 R           emissions of polid in hour 0000-0100 (ug/hr)          Y
adjest02 R           emissions of polid in hour 0100-0200 (ug/hr)          Y
adjest03 R           emissions of polid in hour 0200-0300 (ug/hr)          Y
adjest04 R           emissions of polid in hour 0300-0400 (ug/hr)          Y
adjest05 R           emissions of polid in hour 0400-0500 (ug/hr)          Y
adjest06 R           emissions of polid in hour 0500-0600 (ug/hr)          Y
adjest07 R           emissions of polid in hour 0600-0700 (ug/hr)          Y
adjest08 R           emissions of polid in hour 0700-0800 (ug/hr)          Y
adjest09 R           emissions of polid in hour 0800-0900 (ug/hr)          Y
adjest10 R           emissions of polid in hour 0900-1000 (ug/hr)          Y
adjest11 R           emissions of polid in hour 1000-1100 (ug/hr)          Y
adjest12 R           emissions of polid in hour 1100-1200 (ug/hr)          Y
adjest13 R           emissions of polid in hour 1200-1300 (ug/hr)          Y
adjest14 R           emissions of polid in hour 1300-1400 (ug/hr)          Y
adjest15 R           emissions of polid in hour 1400-1500 (ug/hr)          Y
adjest16 R           emissions of polid in hour 1500-1600 (ug/hr)          Y
adjest17 R           emissions of polid in hour 1600-1700 (ug/hr)          Y
adjest18 R           emissions of polid in hour 1700-1800 (ug/hr)          Y
adjest19 R           emissions of polid in hour 1800-1900 (ug/hr)          Y
adjest20 R           emissions of polid in hour 1900-2000 (ug/hr)          Y
adjest21 R           emissions of polid in hour 2000-2100 (ug/hr)          Y
adjest22 R           emissions of polid in hour 2100-2200 (ug/hr)          Y
adjest23 R           emissions of polid in hour 2200-2300 (ug/hr)          Y
adjest24 R           emissions of polid in hour 2300-2400 (ug/hr)          Y
```

The *polid* field will contain one of the following values: ISOP, TERP, OVOC, or MBO.

Section A-4. Day-specific On-road Mobile Source Emissions Estimates

On-road mobile source emissions estimates will be supplied for each day of each episode. Each day of the episode will have one SAS data set from the $EMS_RUN directory in the following format:

$EMS_RUN/gridmvee

```
variable type format description                                          required
-------- ---- ------ ---------------------------------------------------- --------
stid      I          FIPS state code                                      D
cyid      I          FIPS county code                                     D
areatype  I          area type designation for roadway                    D
                     0 - rural roadway
                     1 - urban roadway
factype   I          facility type designation for roadway                D
                     1 - principal arterial, interstate
                     2 - principal arterial, freeways and expressways
                     4 - principal arterial, other
                     5 - collector
                     6 - minor arterial
                     7 - major collector
                     8 - minor collector
                     9 - local roadway
vtype     A  5       vehicle type                                         Y
                     LDA  - passenger cars
                     LDT1 - type one light duty trucks
                     LDT2 - type two light duty trucks
                     MDV  - medium-duty trucks
                     LHD1 - type one light heavy-duty trucks
                     LHD2 - type two light heavy-duty trucks
                     MHD  - medium heavy-duty trucks
                     HHD  - heavy heavy-duty trucks
                     LHV  - line haul vehicles
                     UBUS - urban buses
                     MCY  - motorcycles
                     SBUS - school buses
                     MH   - motor homes
mvprocess A  3       motor vehicle process                                Y
                     EX - exhaust emissions
                     EV - evaporative emissions
techtype  I          technology type                                      Y
                     1 - non-catalyst gasoline
                     2 - catalyst gasoline
                     3 - diesel
icell     I          east-west grid cell number in the domain             Y
jcell     I          north-south grid cell number in the domain           Y
polid     A  5       pollutant ID                                         Y
                     -- TOG for total organic gases
                     -- NOX for oxides of nitrogen
                     -- CO  for carbon monoxide
ee01      R          emissions of polid in hour 0000-0100 (g/hr)          Y
ee02      R          emissions of polid in hour 0100-0200 (g/hr)          Y
ee03      R          emissions of polid in hour 0200-0300 (g/hr)          Y
ee04      R          emissions of polid in hour 0300-0400 (g/hr)          Y
ee05      R          emissions of polid in hour 0400-0500 (g/hr)          Y
ee06      R          emissions of polid in hour 0500-0600 (g/hr)          Y
ee07      R          emissions of polid in hour 0600-0700 (g/hr)          Y
ee08      R          emissions of polid in hour 0700-0800 (g/hr)          Y
ee09      R          emissions of polid in hour 0800-0900 (g/hr)          Y
ee10      R          emissions of polid in hour 0900-1000 (g/hr)          Y
ee11      R          emissions of polid in hour 1000-1100 (g/hr)          Y
ee12      R          emissions of polid in hour 1100-1200 (g/hr)          Y
ee13      R          emissions of polid in hour 1200-1300 (g/hr)          Y
ee14      R          emissions of polid in hour 1300-1400 (g/hr)          Y
ee15      R          emissions of polid in hour 1400-1500 (g/hr)          Y
```

```
ee16      R         emissions of polid in hour 1500-1600 (g/hr)         Y
ee17      R         emissions of polid in hour 1600-1700 (g/hr)         Y
ee18      R         emissions of polid in hour 1700-1800 (g/hr)         Y
ee19      R         emissions of polid in hour 1800-1900 (g/hr)         Y
ee20      R         emissions of polid in hour 1900-2000 (g/hr)         Y
ee21      R         emissions of polid in hour 2000-2100 (g/hr)         Y
ee22      R         emissions of polid in hour 2100-2200 (g/hr)         Y
ee23      R         emissions of polid in hour 2200-2300 (g/hr)         Y
ee24      R         emissions of polid in hour 2300-2400 (g/hr)         Y
```

The *polid* field will contain one of the following values: CO, NOX, or TOG. The definition of TOG is provided in Appendix G.

Section A-7. Definitions of the Various Forms of Organic Gas And Organic Compounds

Nonmethane Hydrocarbons (NMHC)
Nonmethane Organic Gas (NMOG)
Reactive Organic Gas (ROG)
Total Hydrocarbons (THC)
Total Organic Compounds (TOC)
Total Organic Gas (TOG)
Volatile Organic Compounds (VOC)

TOC: By definition, TOC is any carbon containing compound; however, great care must be taken when using this definition because databases like SPECIATE 3.0 define the profiles as TOC but in fact do not contain such compounds as carbon monoxide and carbonic acid yet do contain carbon sulfide.

TOG: TOG is any organic compounds (i.e. carbon containing compounds except carbon monoxide, carbon dioxide, carbonic acid, metallic carbides [e.g. carbon sulfide], and carbonates [e.g. calcium carbonate, ammonium carbonate]).

THC: THC is computed as TOG minus aldehydes where aldehydes are primarily composed of formaldehyde and acetaldehyde.

NMHC: NMHC is defined as THC minus methane.

NMOG: NMOG is defined as TOG minus methane.

VOC: VOC is defined in 40 CFR § 51.10 and in 63 FR 17331, April 19, 1998, as any compound of carbon, excluding carbon monoxide, carbon dioxide, carbonic acid, metallic carbides, and carbonates that participates in atmospheric photochemical reactions. This includes any such organic compound other than the following, which have been determined to have negligible photochemical reactivity:

➢ methane
➢ ethane
➢ methylene chloride (dichloromethane)
➢ 1,1,1-trichloroethane (methyl chloroform)
➢ 1,1,2-trichloro-1,2,2-trifluoroethane (CFC-113)
➢ trichlorofluoromethane (CFC-11)
➢ dichlorodifluoromethane (CFC-12)
➢ chlorodifluoromethane (HCFC-22)
➢ trifluoromethane (HFC-23)
➢ 1,2-dichloro-1,1,2,2-tetrafluoroethane (CFC-114)
➢ chloropentafluoroethane (CFC-115)
➢ 1,1,1-trifluoro-2,2-dichloroethane (HCFC-123)
➢ 1,1,1,2-tetrafluoroethane (HFC-134a)
➢ 1,1-dichloro-1-fluoroethane (HCFC-141b)
➢ 1-chloro-1,1-difluoroethane (HCFC-142b)
➢ 2-chloro-1,1,1,2-tetrafluoroethane (HCFC-124)

- ➤ pentafluoroethane (HFC-125)
- ➤ 1,1,2,2-tetrafluoroethane (HFC-134)
- ➤ 1,1,1-trifluoroethane (HFC-143a)
- ➤ 1,1-difluoroethane (HFC-152a)
- ➤ parachlorobenzotrifluoride (PCBTF)
- ➤ cyclic, branched, or linear completely methylated siloxanes
- ➤ acetone
- ➤ perchloroethylene (tetrachloroethylene)
- ➤ 3,3-dichloro-1,1,1,2,2-pentafluoropropane (HCFC-225ca)
- ➤ 1,3-dichloro-1,1,2,2,3-pentafluoropropane (HCFC-225cb)
- ➤ 1,1,1,2,3,4,4,5,5,5-decafluoropentane (HFC 43-10mee)
- ➤ difluoromethane (HFC-32)
- ➤ ethylfluoride (HFC-161)
- ➤ 1,1,1,3,3,3-hexafluoropropane (HFC-236fa)
- ➤ 1,1,2,2,3-pentafluoropropane (HFC-245ca)
- ➤ 1,1,2,3,3-pentafluoropropane (HFC-245ea)
- ➤ 1,1,1,2,3-pentafluoropropane (HFC-245eb)
- ➤ 1,1,1,3,3-pentafluoropropane (HFC-245fa)
- ➤ 1,1,1,2,3,3-hexafluoropropane (HFC-236ea)
- ➤ 1,1,1,3,3-pentafluorobutane (HFC-365mfc)
- ➤ chlorofluoromethane (HCF-31)
- ➤ 1-chloro-1-fluoroethane (HCFC-151a)
- ➤ 1,2-dichloro-1,1,2-trifluoroethane (HCFC-123a)
- ➤ 1,1,1,2,2,3,3,4,4-nonafluoro-4-methoxy-butane (C F OCH) 4 9 3
- ➤ 2-(difluoromethoxymethyl)-1,1,1,2,3,3,3-heptafluoropropane ((CF) CFCF OCH) 3 2 2 3
- ➤ 1-ethoxy-1,1,2,2,3,3,4,4, 4-nonafluorobutane (C F OC H) 4 9 25
- ➤ 2-(ethoxydifluoromethyl)-1,1,1,2,3,3,3-heptafluoropropane ((CF) CFCF OC H) 3 2 2 2 5
- ➤ methyl acetate and perfluorocarbon compounds which fall into these classes:
  - cylic, branched, or linear, completely fluorinated alkanes
  - cylic, branched, or linear, completely fluorinated ethers with no unsaturations
  - cylic, branched, or linear, completely fluorinated tertiary amines with no
  - unsaturations and
  - sulfur containing perfluorocarbons with no unsaturations and with sulfur bonds only to carbon and fluorine.

In summary, VOC is TOG minus the exempted compounds.

ROG: ROG is VOC plus aldehydes where aldehydes are primarily composed of formaldehyde and acetaldehyde.

Section A-8. Area Source Spatial Surrogates

surrogates.ar

```
variable type format description                                          required
-------- ---- ------ -------------------------------------------------- --------
stid      I  2.     FIPS state code                                      Y
cyid      I  3.     FIPS county code                                     Y
ssc       A  3.     area source spatial surrogate code                   Y
icell     I  4.     east-west grid cell number in the domain             Y
jcell     I  4.     north-south grid cell number in the domain           Y
ratio     R  13.    fractional value of ssc in stid, cyid, icell,        Y
                    and jcell
```

Section A-9. Area Source Category-to-Area Source Spatial Surrogates Cross Reference

asct2surrogates.ar

```
variable type format description                                         required
-------- ---- ------ -------------------------------------------------- --------
ssc       A  3.     area source spatial surrogate code                  Y
asct      A  15.    area source category                                Y
```

Section A-10. VOC-to-TOG Conversion Factors By Source Category

voc2tog.sp

```
variable type format description                                             required
-------- ---- ------ -------------------------------------------------- --------
asct or   A  15.    area source category or source classification code   Y
scc
factor    R  13.    factor applied to convert VOC emissions to TOG       Y
                    emissions
```

Section A-11. Chemical Mechanism-specific Hydrocarbon Speciation Profiles

profile.sp

```
variable type format description                                        required
-------- ---- ------ ---------------------------------------------------- --------
inprf     A    5.    hydrocarbon speciation profile                        Y
mech      A    10.   chemical mechanism identifier                         Y
                     CBIV  – cabon bond version four
                     SAPRC – SAPRC version 1999
mdlspab   A    5.    chemical mechanism model species abbreviation         Y
factor    R    13.   factor to convert emissions of TOG to emissions       Y
                     of mdlspab
```

Note, emissions of NO_X, in grams, are converted to NO and NO_2, in moles, by multiplying NO_X emissions by 0.01957 and 0.00217 respectively. Emissions of CO, in grams, are converted to CO, in moles, by multiplying CO emissions by 0.03571.

Section A-12. Source Category-to-Hydrocarbon Speciation Profile Cross Reference

scc2profile.sp

```
variable type format description                                          required
-------- ---- ------ -------------------------------------------------- --------
asct or   A    15.   area source category or source classification code   Y
scc
inprf     A    5.    hydrocarbon speciation profile                       Y
```